AFOSR-TR- 95 0101

Fundamental limitations and Capabilities of Robust Identification and Adaptive Control
Final Report

M.A. Dahleh and J.N. Tsitsiklis

Department of Electrical Engineering and Computer Science

Massachusetts Institute of Technology

Cambridge, MA 02139

January 17, 1995

Abstract

In this report, we summarize the major results obtained under grant AFOSR-91-0368.

Accesion For	
NTIS CRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

Contents

1 Introduction

In the past decade, a powerful theory for designing robust control systems has emerged. Starting with a model, and a description of the uncertainty (structured, parametric etc.), a controller can be designed to meet a variety of performance specifications. This development, however, has not been accompanied by a parallel development in system identification methods by which a plant model *and* a description of uncertainty is provided. In an attempt to bridge this gap, a new area of research in robust identification has emerged in the last few years. This research is motivated (in part) by the following:

1. It is evident that a "good" controller cannot be designed based entirely on a model *without* a description of plant uncertainty [15, 16, 8]. Current identification schemes do not provide information about plant uncertainty that is *usable* by current robust control techniques [36].

2. The failure of most adaptive systems is a consequence of the failure of the identification scheme within the adaptive controller. This failure can be described either in terms of parameter convergence (a traditional and possibly inappropriate description), or in terms of plant uncertainty [2, 44].

3. Much of the research done up to now on system identification has assumed that the noise process is stochastic, e.g., filtered white noise, with stationarity being an important side assumption. A lot of attention has been paid to showing convergence, as well as to deriving bounds on confidence intervals, all *asymptotically.* Not much effort was put into problems with finite data and possibly nonstationary noise.

4. The status of spectral estimation remains as in Jenkins and Watts [27]. For nonstationary noise, much of that theory does not yield satisfactory results.

5. There was very little understanding of the fundamental limitations of system identification in the presence of different classes of noise, and when the objective is to reduce the plant uncertainty given only finite data.

6. The limitations of controller design when only finite corrupted data is available are not well understood. In that sense, the available tools from robust control are not well connected with experiments, and the assumptions underlying the existing paradigms may be somewhat unrealistic.

As a result, there has been increasing interest, among the control and identification communities, in the problem of identifying plants for control purposes. This generally means that the identified model should approximate the plant as it operates on a rich class of signals, namely signals with bounded norm, since this allows for the immediate use of robust control tools for designing controllers. This problem is of special importance when the data are corrupted with bounded noise. The case where the objective is to optimize prediction for a fixed input was analyzed by many researchers [18, 34, 37, 38, 39, 42]. The problem is more interesting when the objective is to approximate the original system as an operator, a problem extensively discussed in [55]. For linear time invariant plants, such approximation can be achieved by uniformly approximating the frequency response (in the $\mathcal{H}_\infty$-norm) or the impulse response (in the ℓ_1 norm). In $\mathcal{H}_\infty$ identification, it was shown that robustly convergent algorithms can be furnished, when the available data is in the form of a corrupted frequency response, at a set of points dense on the unit circle [22, 23, 24, 20, 21]. When the topology is induced by the ℓ_1 norm, a complete study of asymptotic identification was given in our past work [52, 53, 54] for arbitrary inputs, and the question of optimal input design was addressed as well. Related work on this problem was also reported in [19, 26, 31, 35, 40, 41, 48].

Another issue of importance in the context of worst-case identification is *complexity*. It turns out that it is generally much harder to devise experiments that can guarantee small worst-case errors in the presence of bounded noise. This problem has been extensively analyzed in our work [11] and elsewhere [43, 33].

It is important to caution at this point regarding the meaning of "worst-case" errors. that the terminology "worst-case" does not mean that one can furnish guarantees on the worst-case error with respect to the actual plant. Clearly, any result we obtain is a function of prior assumptions (which are not verifiable in general), and thus the results hold only when these assumptions are valid. This is no different from the traditional stochastic approach for system identification. One cannot derive guarantees about the actual plant, from only finite data, without additional assumptions about the set of possible plants, and any methodology will be subject to this limitation.

Even with this recent development, system identification and robust control remain separate fields. The estimates of uncertainty obtained from the above methods tend to be quite conservative, which renders them useless for robust control methods. A framework unify-

ing the controller design problem has to be iterative in nature, and robust control methods should play a role in the selection of experiments for the next iteration. In this sense, the hypothesized model structures should include a description of the uncertainty (that will not be identified). Once such a model is described, a controller can be designed based on this description. The signals used to test this controller should provide further useful data to tune this model further and obtain better performance at the next iteration. Needless to say, a computable theory of this kind is still not available. Iterative identification/control methods have already been discussed in the literature (see [1, 30, 46, 56] for example). However current approaches are based on simply attempting to identify the system in closed loop, refining the control design and the identified model as the iterations proceed. As such these methods merely aim towards a particular closed loop model (for a specific controller). Even though these methods depart from the traditional system identification approach, they still do not provide a framework in which information from a previous iteration reduce the plant uncertainty for the next iteration. What is lacking is a general and systematic means to exploit powerful robust control design and set membership identification techniques, and hence provide an identification and control design methodology with firm performance guarantees.

Our research addresses the general controller design problem starting from finite corrupted data and some prior information. On one hand, we will study the identification problem in the presence of deterministic/stochastic noise, and study the fundamental limitations and capabilities of identification in such a setup. In particular, we will study the problem of translating this coarse description of the experimental setup, into a precise description of a plant and uncertainty. On the other hand, we will develop robust control techniques to handle the most general robust performance problem. We will show how these can be integrated into one framework in which identification and control are done in an iterative fashion. While this will provide a procedure for a systematic design, it is still far from feasible with current methods, and our research will concentrate on providing the tools for implementing it.

1.1 Summary of Past Accomplishments

Our past research has concentrated on developing a theoretical foundation for system identification in the presence of deterministic noise. In particular, the work of Tse *et al* [10, 52, 53, 54] allows for the analysis of large classes of systems, including nonlinear fading memory

systems. The study is done in two steps. The first step is concerned with obtaining tight upper and lower bounds on the optimal achievable error, for a given fixed experiment. The second step is to study these bounds and characterize the inputs that will minimize them. The upper and lower bounds are obtained under some mild topological assumptions on the model set, and for any fixed experiment, through the diameter of the worst-case uncertainty set, a concept borrowed from Information Based Complexity [49, 50].

Using this formulation, we have studied in detail several model sets containing linear time invariant stable systems. We also analyzed the sample complexity in the case of unknown bounded noise.

Our research in robust control has concentrated on developing computational methods for solving the ℓ_1 robust control problem. These methods can be extended to incorporate additional frequency-domain and time-domain constraints that are not directly captured by standard theory. The methods provide bounds on the optimal achievable performance and give information about the structure of the optimal controller. This work has formed the basis of some software tools that we have developed for designing control systems in the presence of mixed objectives. Finally, in a related effort, some major open problems in robust control have been addressed using the theory of computational complexity.

A last area of research has dealt with the foundations of learning theory, as developed by computer scientists and statisticians, with the objective of linking it to the basic problems of learning that arise in control theory.

Part of our effort has been channelled towards education. In that regard, we have written a textbook explaining the current robust control paradigm emphasizing computations. The book is titled: Control of Uncertain Systems: A Linear Programming approach (by Dahleh and Diaz-Bobillo). In the book, we present a unifying theory for robust control that is quite accessible to engineers at all levels. This will help in bridging the existing gap between theory and applications.

2 Details of Past Research

2.1 Robust Identification

We consider a framework for system identification which is meant to provide not only a nominal model for an unknown plant, but also some hard guarantees on the distance of the

true model from the nominal. For clarity of exposition, the discussion that follows is based on a concrete set of assumptions. However, the framework is more general and we discuss alternative settings as we proceed.

We start with a model set $\mathcal{M}$ which is meant to capture any prior information we might have on the unknown system to be identified. For example, $\mathcal{M}$ could be the set of all stable linear time invariant systems, or the set of all LTI systems with a finite impulse response of length N. Let $\mathcal{U}$ be the set of all inputs $u(\cdot)$ such that $|u(t)| \leq 1$ for all t. Finally, let $\mathcal{D}$ be the set of all output disturbances $d(\cdot)$ that satisfy $|d(t)| \leq \delta$ for all t, where δ is a given constant. We then consider the following sequence of events. We choose some input function $u \in \mathcal{U}$ and apply it to the unknown system $h \in \mathcal{M}$. We observe a noise–corrupted output of the system, of the form

$$y = u * h + d, \tag{1}$$

where $d \in \mathcal{D}$. Based on the observation y and our knowledge of u, $\mathcal{M}$, and $\mathcal{D}$, we can form the *uncertainty set* $\Omega(y)$ which is the set of all models that are possible, given the information that we have:

$$\Omega(y) = \{h \in \mathcal{M} \mid \exists d \in \mathcal{D} \text{ such that } y = u * h + d\}. \tag{2}$$

We might choose an element $\hat{h}$ of $\Omega(y)$ and call it the estimate of h. The worst case error is

$$E(y) = \sup\{\|h - \hat{h}\| \mid h \in \Omega(y)\},$$

where $\|\cdot\|$ is a norm on the set of all plants. In fact, no matter how we choose $\hat{h}$, we have

$$\frac{1}{2}\text{diam}(\Omega(y)) \leq E(u) \leq \text{diam}(\Omega(y)).$$

Thus, instead of focusing on any particular estimate, we might concentrate on the diameter of the uncertainty set $\Omega(y)$ and view it as a measure of the identification accuracy we have achieved.

In earlier research [52, 53], we have provided a conceptual foundation for the above outlined approach. We have proved that in the limit of very long experiments, the best achievable diameter $\inf_{u \in \mathcal{U}} E(u)$ is either equal to 2δ or it is infinite. Which of the two will be the case depends on the underlying model set, that is, on the amount of prior available information. This allows us to say that some model sets are *learnable* and some are not, depending on the value of $\inf_{u \in \mathcal{U}} E(u)$. We have also shown that a model set is learnable if

and only if, under our experimental setup, we can distinguish between stable and unstable plants.

In another study [11, 51], we focused on worst–case identification, under the ℓ_1 norm error criterion, of plants with a finite impulse response. Although, this is a learnable model set (the worst–case error can be made as small as 2δ), we have proved that the experiment length must be an exponential function of the length of the impulse response, even if we are willing to settle for an error which is within a constant factor of δ. This results suggests that the standard assumptions used in worst–case identification are too conservative to be practical, and that some probabilistic aspects should be introduced.

Our most recent work in this area [5], has used an alternative and possibly more realistic model of the noise sequence d, commonly referred to as "deterministic white noise." With this model, the set $\mathcal{D}$ of admissible disturbances is constrained further by requiring the sample autocorrelation of the disturbance sequence d to be low, which provides a deterministic counterpart of white noise. Our work has provided upper and lower bounds on the worst–case diamater of the uncertainty set, as well as exponential lower bounds on the length of the experiments required to obtain a small enough diameter.

2.2 Robust Control

We summarize below our research accomplishments in the area of robust control.

1. Computation of ℓ_1 Optimal Solutions:

The contributions in this regard are marked by the introduction of the Delay Augmentation Algorithm for solving nonsquare problems (e.g., problems with more regulated variables than actuators) [13, 14]. This algorithm is based on squaring the system by introducing fictitious delayed inputs and outputs. The problem is solved iteratively as the number of delays increase. At each iteration, a square ℓ_1 problem is solved (the solution of which is known exactly). The main features of this algorithm are that: (1) at each iteration it gives upper and lower bounds of the optimal objective function which are convergent; (2) it provides information about the structure of the controller; (3) it does not cause order inflation (it is not based on FIR approximations); (4) it involves solving one linear program iteratively. In many cases, the exact solution for nonsquare problems is provided.

For implementation purposes, all computations are performed using matrix algebra, often exploiting the structure of Toeplitz matrices resulting from convolution operators. An

example of that is the development of methods for computing directions of zeros with multiplicity using Toeplitz matrix manipulation, without ever computing the Smith-McMillan Decomposition.

A major part of this research that parallels our research in computation has been the development of software. Using this software, we have studied a variety of benchmark problems (e.g., the X29 Aircraft, an earth-observing system (EOS-AM), a flexible beam, a high purity distillation column). The following are the main new features of the software.

1. Using *Delay Augmentation* as the main core for nonsquare problems.
2. Characterizing feasible subspaces by zeros. The computations involve lower triangular block-Toeplitz matrices.
3. All necessary computations are in state-space.
4. Optimization involves solving linear programs.

2. Robustness Analysis and Synthesis

This is concerned with the development of a computational theory to address directly uncertain plants. The uncertainty is structured in nature, possibly time varying, but nonparametric. In this regard, we have built on the results in [9, ?] to come up with simple conditions for ℓ_∞ robust analysis in the presence of structured uncertainty [8]. This can be readily generalized to MIMO perturbation blocks. The conditions are stated in terms of the spectral radius of a matrix constructed from computing the ℓ_1 norms of certain closed loop maps. We have also analyzed the case of time-invariant perturbations when ℓ_∞ stability is required and have shown that the natural conditions are in the frequency domain (coincide with the standard μ results).

Since the spectral radius of a positive matrix can be computed by minimizing a scaled ℓ_1 norm, synthesis for structured uncertainty problems involves iterations between solving an ℓ_1 problem and finding optimal scales for the uncertainty. We have analyzed this algorithm in detail, and have shown its limitations. We have also proposed an alternative algorithm based on sensitivity analysis of the linear programming solution of the ℓ_1 problem [45].

Finally, we have looked at some of the basic problems of robust control, using the tools of the theory of computational complexity. For example, suppose that we are given interval matrices A, B, and C (that is, matrices with each entry being a range of possible values). A most basic problem in robust control is to find a feedback gain matrix K such that

$A + BKC$ is stable for all possible matrices A, B, C whose entries are within the allowed ranges. We have shown that a version of this problem, as well as some related problems in decentralized control and simultaneous stabilization, are NP–hard [4]. This means that under the prevailing conjectures in computational complexity theory, these problems are not efficiently solvable. Results of this type are useful in determining the fundamental limits of what types of solutions to such problems are possible, and also determine what kind of research can be meaningfully pursued.

3. Writing a Book on Robust Control

The book titled: *Control of Uncertain Systems: A Linear Programming Approach* written by Dahleh and Diaz-Bobillo presents a unified treatment of the theory of robust control design with emphasis on computational methods. It can serve as a starting point for researchers in the field as well as a textbook for a graduate class in control. In our opinion, this is the only book available that gives a comprehensive treatment of $\mathcal{H}_2$, $\mathcal{H}_\infty$ and ℓ_1 methods integrated in a *robust performance* framework, with emphasis on computations.

2.3 Identification for Controller Design

The traditional route to controller design has been to first perform some system identification, so as to abstract a mathematical model from the physical process. This model is then used as the basis for the design of the controller. In some instances the controller may not have the desired properties when implemented on the actual system. In these cases, one has to go back and alter the design in some fashion. However it is often not clear whether the fault lies in the controller design process, the system identification procedure, or stems from the fact that one has not taken enough data to properly identify the plant, or has set performance specifications that are simply too stringent to be met.

The problems arise from the fact that this traditional route is rather ad-hoc, so that when it fails one does not know where to lay the blame. We would like to develop a framework for addressing these issues in a systematic and quantifiable fashion. In order to do so, we first note that our ultimate goal is to design a controller which meets the required performance specifications *on the actual plant.* With this observation we see that there is no necessity to artificially split this design process into an identification procedure, and a control law design. Moreover we believe that by so doing one throws away a lot of potentially useful information.

1. Problem Definition

The basic control problem we consider may be stated as follows: *Given some prior information about the process and a set of finite data, design a feedback controller that meets the given performance specifications.* We propose to develop a framework for addressing this problem by considering an integrated system identification and design process. The resulting procedure should allow us to incorporate an array of system identification and controller design methodologies, so that we may adapt our methods to make use of the latest tools. We will require that the design procedure be systematic, so that at each stage the next course of action is clear, and the procedure terminates with a successful controller design or the conclusion that the performance specifications cannot be met (subject to our *prejudice*).

3. An Iterative Formulation

Rather than assuming that the prior information is true, it is more natural to think of it as a parametrization of model structures from which we desire to explain the data, i.e., a description of our *prejudice*. In this sense, prior information can itself be invalidated by the data. This distinction is crucial since such information is generally derived from simplified models of the process, and hence is not verifiable. Once a set of finite data is acquired, a set of models that are consistent with the data and the model structure parametrization (prior information) is defined. This set contains all models that are not falsified by the data. Roughly speaking, system identification picks a most powerful unfalsified model where most powerful is defined depending on the objective in mind. In this case it is finding a controller that delivers a given performance level. We also note that the process of finding such a model, and a controller, is iterative in nature as more sets of data are acquired.

It is evident that any iterative scheme will generally be based on reducing the set of unfalsified plants until a controller based on the remaining elements can deliver the performance when connected with the actual process, or a decision is made to enlarge the parametrized set of models and/or change the performance requirement. We propose a general scheme that is based on efficiently eliminating models from the set of unfalsified models. Of course, the acquisition of more data systematically reduces this set, although the efficiency of this depends on the data set itself. On the other hand, an unfalsified model is invalidated if there exists a controller that delivers the required performance for this model and the same controller does not meet the performance with the actual process. Given our prejudice, this

model is unacceptable. Finally, an unfalsified model to which no controller can be designed to meet the performance specifications is discarded. In this way, if all models are eliminated, we conclude that the performance cannot be met, given our prejudice. Below, an iterative scheme based on this idea is proposed [7]. This scheme is well defined only if we assume that the required performance of any controller connected with the real process can be tested by using a finite number of experiments. Of course in practice these are the only performance requirements that we can ever verify.

1. Pick a model structure parametrization.

2. Collect a set of data, and define the set of unfalsified plants.

3. Find a "large" subset of models for which the design procedure yields a controller that delivers the required performance for all models in this set. If no such set exists, go back to (1) and adjust the model structure and/or the performance objective.

4. Test the controller on the real system. If the controller meets the performance, then stop. If not, then the above subset is invalidated.

5. Use the data acquired from testing the performance, as well as other sets of data, in order to invalidate additional plants.

6. Go to (3).

This scheme defines both an inner and outer loop. Within the inner loop, the performance requirement and the model structure parametrization are fixed, and the acquisition of data, as well as the design of controllers for subsets of the set of unfalsified models, continue to reduce this set until a controller is found, or a decision that the performance cannot be met is made. We then iterate the outer loop. By eliminating large subsets in step (3), the inner loop converges to a decision much faster.

The process of elimination requires the availability of methods for designing robust controllers for subsets of the set of unfalsified models. It is assumed that for a given subset, a decision can be made as to whether or not a controller that meets the performance exists. If the parametrization of models and the performance objective are such that no exact methods exist, one may use tests based on the existing design methods, as conservative as they may be. The lesser the conservatism of the methods in robust control, the lesser the bias of the above iterative scheme will be.

The step of testing a given controller on the real system generates more sets of data that can be used to invalidate more models, within the inner loop of the above scheme [32]. We may have the ability to conduct more experiments, in which case they have to be devised in such a way that they have sufficient information to invalidate more unfalsified models. The design of such experiments is one of the research directions to be addressed.

Note that the system identification and control design procedures are not distinct in this framework, but intimately connected. The process of refining both the model and the control design takes place concurrently. This allows us to exploit the full power of new set membership identification procedures, and robust control design techniques, to work with *sets of plants* for both modeling and design. The potential of this scheme for improving over existing techniques arises largely from exploiting the connections between these two fields.

2.4 Learning Theory

Starting with the seminal work of Vapnik and Valiant, there has been a surge of activity in computational learning theory, whose objective is to characterize what can be learned, and how much information is required for effective learning to take place. Although this theory had not been linked to system identification and control theory, the basic questions raised are similar in both areas.

An important factor in system identification is that the experimenter can choose what inputs to apply. In an abstract setting, this amounts to "active learning" whereby the experimenter has latitude as to the type of information to be obtained. Our results [28] have established that whatever is learnable by active learning is also also learnable under "passive" learning, but active learning reduces the amount of experimentation required. In addition, our work has highlighted the fundamental role of metric entropy, which leads to some intriguing possibilities of establishing a connection with the control–theoretic work of Zames [55].

Finally, in other work [29], we have extended the traditional model of computational learning theory ("PAC learning") by introducing and studying the notion of "generalized samples." Besides the applications in image analysis that were discussed in [29], such extensions of the traditional model may prove useful in bridging the gap with the discipline of system identification.

3 Industrial Interactions

3.1 Professor Dahleh's Industrial Interactions

Professor Dahleh has been involved in the development of a design methodology based on the ℓ_1 theory (an effort supported in part by AFOSR), which has been fully explained in the recent book [6]. To make this methodology accessible to industry, he has been involved in developing software based on matlab for synthesis of controllers for plants with uncertainty. This software has interactive features by which a design can be altered by graphically changing the various responses or frequency plots of the system. Version of this software are currently available through ftp.

Professor Dahleh has been working very closely with C.S. Draper Laboratory in the areas of robust control and system identification. In the area of robust control, he has educated several groups about the various robust control tools, as well as the ℓ_1 software. The latter is now a standard tool used by all the engineers working in the control division. Most recently, Professor Dahleh and his students have been involved in the attitude control problem of the Earth-Observing Satellite. (The ℓ_1 methodology is the right formulation for this problem for a variety of reasons: The first has to do with the specifications being in the time domain in terms of limits of allowable deviation of attitude angles. Secondly, the constraints are in terms of saturating gyros (due to accumulated momentum). And finally, the class of plant uncertainty includes nonlinearities as well as time variation.) Professor Dahleh and his students have done designs using both H_∞ and ℓ_1 and shown that ℓ_1 can exhibit the limits and tradeoffs of the design in a much more systematic fashion. In fact, the H_∞ designs had responses that are quite inferior to the design they exhibited.

In the system identification area, Professor Dahleh has been supervising the implementation of the recently developed iterative control and identification scheme (which is also developed under our AFOSR grant). The objective is to develop a CAD environment by which controllers can be designed directly from Data. The controllers are then changed as more testing Data is acquired. C.S. Draper Laboratoy plans to have such an environment available as a tool for designing control systems.

Professor Dahleh has also been working very closely with FIAT research center, and recently with Ford (a starting effort), on the design of active suspensions. The suitability of the ℓ_1 problem is also clear for this application.Professor Dahleh has educated engineers to help them use the software to design such systems. Also, he has recently done a complete

case study on this problem exhibiting the exact tradeoffs between the specifications and the constraints. Ford is also interested in developing capabilities for iterative identification and control for direct use in their test environment (they have a complete computational facility inside the test cars, to update the controller design from the testing data). The setup Professor Dahleh has proposed as a result of his research in the system identification area appears to be quite attractive.

Recently, Professor Dahleh in collaboration with other faculty at MIT has acquired a contract from Siemens to develop capabilities for identification of nonlinear systems and to develop an iterative identifcation/control environment. Developing software is a major part of that effort.

Professor Dahleh has also been in close contact with government Labs. In particular, he has very close ties with Dr. Ridgley at Wright Patterson. Dr. Ridgely has been involved in studying mixed optimization problems and recently has been studying the ℓ_1 design methodology. Professor Dahleh has provided him with draft copies of the book [1] as well as access to the software. Dr. Ridgely taught a course from [6], and many of his students are now well versed with the current robust control theory, including ℓ_1. It is intended to push this collaboration further and educate several engineers at WPAFB to use the ℓ_1 software, which will be available very soon. This will be accomplished by giving short courses, and demonstrations of the software on site. In addition, Professor Dahleh has developed very close ties with Dr. Coleman in one of the Army Labs (ARDEC).

Finally, Several Engineers have already started using Professor Dahleh's software at Hughes. Also, several engineers are investigating using the software for noise cancellation application (vibration suppression) at BBN. Professor Dahleh has also made initial contacts with several companies interested in control and identification (e.g. Speyer which is interested in semiconductor devices, elgin Bailey and Bailey controls which are interested in the problem of integration of several control systems). In addition, Professor Dahleh's work in system identification has had a large impact on the space lab at MIT. The objective of the experiments is to study the modeling problem for the purpose of control. Professor Dahleh has served on several thesis committees and was quite influential in guiding the research in that discipline.

3.2 Professor Tsitsiklis Industrial Interactions

Although Prof. Tsitsiklis' industrial interactions are not directly linked to the core subjects of the research performed under this grant, there have been extensive such interactions that fall within the broader themes of systems and control theory.

For example, Prof. Tsitsiklis has been working together with the C. S. Draper Laboratory, towards the development of hierarchical control architectures for the planning and operation of advanced train control systems. This work involves the application of the decomposition methods described in [3], to the large scale planning and scheduling problems faced by railroad companies.

One of the directions towards which Prof. Tsitsiklis' research is moving is the application of function approximation methods in the computation of the optimal cost–to–go function of dynamic programming in order to bypass the curse of dimensionality that plagues nonlinear control problems. His research in this area is already being transferred to the commercial sector, by a number of companies dealing with scheduling, resource allocation, and logistices problems. This line of research should be of interest to the Air Force on several counts. First, because the Air Force is faced with several challenging logistics problems; second, because with the accumulation of experience, we expect to be able to solve in the near future, nontrivial problems involving the control of complex dynamical systems.

In another effort, Prof. Tsitsiklis and two more M.I.T. faculty have launched a research program with the Groupe Schneider and Square D whose goal is to "reengineer" the basic architectures used in industrial automation and to envision the technology that will take the place of Programmable Logic Controllers (PLCs). This research taking place in the context of frequent site visits and close technical interaction with Groupe Schneider engineers.

Finally, Prof. Tsitsiklis has initiated a collaboration with faculty in the M.I.T. department of chemical engineering whose goal is to apply neural network techniques for the analysis of pharmaceutical process data, with the aim of identifying "signatures" that can be used for early prediction of the performance of a batch as well as of identifying control variables that can be manipulated so as to enhance performance.

In conclusion, the work of Professors Dahleh and Tsitsiklis has been coupled directly with several industrial activities. These activities have been quite extensive and have shaped the direction of our present research directions. In addition, both are working closely with Dr.

Gunter Stein who has been instrumental in shaping the research effort in the control area at MIT.

4 Educational Impact

This research has supported three excellent Masters theses and one Ph.D thesis in the area of system identification. The first S.M. thesis was by David Tse, in which the problem of worst-case identification in the presence of bounded noise was completely covered. The second S.M. thesis was by Theodore Theodosopoulos in which the problem of sample complexity of worst-case identification was formulated and solved. The third S.M. thesis was by Ian Chen, which generalized Tse's results for bounded noise with low correlations. The Ph.D thesis was partially supported by this grant in which the problem of iterative identification and control was formulated and discussed.

In the area of robust control, this grant supported in part one major Ph.D thesis by Ignacio Diaz-Bobillo, which contained a major development of the ℓ_1 theory. The software development was also a result of this work. Also, this research supported in part a S.M. thesis that applied the software to the earth-observing system (EOS-AM). The latter thesis demonstrated the power of this ℓ_1 theory in achieving high precision in pointing applications.

This grant also supported in part the work on the book [6]. This book is now used in several universities and industrial laboratories.

References

[1] B. Anderson and R. Kosut, "Adaptive Robust Control: On–Line Learning", *Proc. 1991 Conference on Decision and Control.* Brighton, England, December 1991.

[2] K.J. Astrom, "Analysis of Rohr's counterexample to adaptive control," *Proc. 22nd IEEE conf. on Dec. & Cont.*, San Antonio, TX, U.S.A.,1983.

[3] D. P. Bertsekas, J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Prentice Hall, 1989.

[4] V. Blondel and J. N. Tsitsiklis, "NP–hardness of some linear control design problems", submitted to the *SIAM J. on Control and Optimization*; technical report LIDS-P-2257, August 1994.

[5] K. H. Chen, "System Identification under low-correlated noise", Masters Thesis, M.I.T. January 1995.

[6] M.A. Dahleh and I. Diaz-Bobillo. *Control of Uncertain Systems: A Linear Programming Approach*, Prentice-Hall, N.J., 1995.

[7] M.A. Dahleh and J.C. Doyle. "From Data to Control," Springer Verlag, special edition on identification for robust control, edited by R. Smith and M. Dahleh.

[8] M.A. Dahleh and M.H. Khammash, "Controller Design in the Presence of Structured Uncertainty," to appear in *Automatica* special issue on robust control.

[9] M.A. Dahleh and Y. Ohta. "A necessary and sufficient condition for robust BIBO stability," *Systems and Control Letters*, Vol. 11, pp. 271-275, 1988.

[10] M.A. Dahleh, E.D. Sontag, D.N. Tse and J.N. Tsitsiklis, "Worst-Case identification of nonlinear fading memory systems," *ACC Proc.*, 1992, to appear in Automatica.

[11] M.A. Dahleh, T. Theodosopoulos, and J.N. Tsitsiklis, "The sample complexity of worst-case identification of F.I.R. Linear systems," *System and Control. letters*, Vol 20, No. 3, pp. 157-166, March 1993.

[12] M.A. Dahleh, T. Theodosopoulos, and J.N. Tsitsiklis, "The sample complexity of worst-case identification of F.I.R. Linear systems," *Proc. of the CDC*, December 1993.

[13] I. Diaz-Bobillo and M.A. Dahleh, "State Feedback ℓ_1 Optimal Controllers Can be Dynamic," *System and Control. letters*, Vol 19, No. 2, pp. 87-93, August 1992.

[14] I. Diaz-Bobillo and M.A. Dahleh. "Minimization of the maximum peak-to-peak gain: The General multiblock problem," *IEEE Trans. Automat. Contr.*, Vol 38, No. 10, pp. 1459-1482, October 1993.

[15] J. C. Doyle, "Analysis of feedback systems with structured uncertainty," *IEEE Proceedings* 129, 242-250,1982.

[16] J.C. Doyle and G. Stein. "Multivariable feedback design: concepts for a classical/modern synthesis," *IEEE-Trans. A-C*, Vol. 26, No. 1, February 1981.

[17] N. Elia, M.A. Dahleh and I.J. Diaz-Bobillo. "Controller design via linear programming," *Proc ACC*, San Francisco, CA, June 1993, pp. 2165-2169.

[18] E. Fogel and Y. F. Huang, " On the value of information in system identification–bounded noise case", *Automatica*, vol.18, no.2, pp.229-238, 1982.

[19] G.C. Goodwin. "Experiment design in system identification" in *Encyclopedia of Systems and Control* (M. Singh, Ed.), Pergamon Press, 1987.

[20] G. Gu, P.P. Khargonekar and Y. Li, "Robust convergence of two sta ge nonlinear algorithms for identification in $\mathcal{H}_\infty$, *Systems and Control Let ters*, Vol 18, No. 4, April 1992.

[21] G. Gu and P.P. Khargonekar, " Linear and nonlinear algorithms for identification in $\mathcal{H}_\infty$ with error bounds, *IEEE Trans. A-C*, Vol 37, No. 7, July 1992.

[22] A.J. Helmicki, C.A. Jacobson and C.N. Nett, "Identification in H_∞: A robust convergent nonlinear algorithm", Proceedings of the 1989 International Symposium on the Mathematical Theory of Networks and System, 1989.

[23] A.J. Helmicki, C.A. Jacobson and C.N. Nett, "Identification in H_∞: Linear Algorithms", Proceedings of the 1990 American Control Conference, pp 2418-2423.

[24] A.J. Helmicki, C.A. Jacobson and C.N. Nett, "Control-oriented System Identification: A Worst-case/deterministic Approach in H_∞'," *IEEE Trans. A-C*, Vol 36, No. 10, October 1991.

[25] H. Hjalmarsson and L. Ljung, "A discussion of unknown-but-bounded disturbances in system identification," *Proc. of the CDC*, December 1993.

[26] C.A. Jacobson and C.N. Nett, "Worst-case system identification in ℓ_1: Optimal algorithms and error bounds," in *Proc. of the 1991 American Control Conference*, June 1991.

[27] G.M. Jenkins and D.G. Watts. *Spectral analysis and its applications*, Holden-Day, 1968.

[28] S.R. Kulkarni, S.K. Mitter and J.N. Tsitsiklis, "Active Learning Using Arbitrary Binary Valued Queries", *Machine Learning*, Vol. 11, 1993, pp. 23–35.

[29] S.R. Kulkarni, S.K. Mitter, J.N. Tsitsiklis, and O. Zeitouni, "PAC Learning with Generalized Samples and an Application to Stochastic Geometry", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 9, September 1993.

[30] W. Lee and B. Anderson and R. Kosut and I. Mareels, "On Adaptive Robust Control and Control-Relevent System Identification", *Proc. 1992 American Control Conference, Chicago, IL*, June 1992.

[31] J.M. Krause. *Robust identification and adaptation*, Ph.D Thesis, University of Minnesota, 1987.

[32] M.M. Livstone, M.A. Dahleh and J.A. Farrell. "A framework for robust control based model invalidation, submitted to *ACC*.

[33] L. Lin, L. Wang and G. Zames, "Uncertainty principle and identification n-Widths for LTI and slowly varying systems, " *ACC*, Chicago, IL, 1992.

[34] R. Lozano-Leal and R. Ortega, "Reformulation of the parameter identification problem for systems with bounded disturbances", *Automatica*, vol.23, no.2, pp.247-251, 1987.

[35] M.K. Lau, R.L. Kosut, S. Boyd, "Parameter Set Estimation of Systems with Uncertain Nonparametric Dynamics and Disturbances", Proceedings of the 29th Conference on Decision and Control, pp. 3162-3167, 1990.

[36] L. Ljung. *System identification: theory for the user*, Prentice-Hall, Inc, NJ, 1987.

[37] M. Milanese and G. Belforte, "Estimation theory and uncertainty intervals evaluation in the presence of unknown but bounded errors: Linear families of models and estimators", *IEEE Trans. Automatic Control*, AC-27, pp.408-414, 1982.

[38] M. Milanese and R. Tempo, "Optimal algorithm theory for robust estimation and prediction", *IEEE Trans. Automatic Control*, AC-30, pp. 730-738, 1985.

[39] M. Milanese, "Estimation theory and prediction in the presence of unknown and bounded uncertainty: a survey", in *Robustness in Identification and Control*, M. Milanese. R. Tempo, A. Vicino Eds, Plenum Press, 1989.

[40] P.M. Makila, "Robust Identification and Galois Sequences", Technical Report 91-1, Process Control Laboratory, Swedish University of Abo, January, 1991.

[41] P.M. Makila and J.R. Partington, "Robust Approximation and Identification in H_∞", Proc. 1991 American Control Conference, June, 1991.

[42] J.P. Norton, "Identification and application of bounded-parameter models", *Automatica*, vol.23, no.4, pp.497-507, 1987.

[43] K. Poolla and A. Tikku, "On the time complexity of worst–case system identification," to appear in IEEE Trans. AC.

[44] C.E. Rohrs, L. Valavani, M. Athans and G. Stein. "Robustness of continuous-time adaptive control algorithms in the presence of unmodeled dynamics," *IEEE Trans. Automat. Contr.*, vol. AC-30, no. 9, pp. 881-889, September 1985.

[45] V. Saligrama. "Properties of ℓ_1 optimal solutions," MIT S.M. Thesis, February 1993.

[46] R.J. Schrama, "Accurate identification for control: the necessity of an iterative scheme," Selected topics in Identification modeling and control, vol 4, 1992.

[47] R. Schrama and P. Van den Hof, An Iterative Scheme for Identification and Control Design Based on Coprime Factorizations", *Proc. 1992 American Control Conference*, Chicago, IL, June 1992.

[48] R. Smith and M. Dahleh. *The Modeling of Uncertainty in Control Systems*, Springer-Verlag, 1994.

[49] J.F. Traub and H. Wozniakowski, *A General Theory of Optimal Algorithms*, Academic Press, New York, 1980.

[50] J.F. Traub, G. Wasilkowski and H. Wazniakowski, *Information-Based Complexity*, Academic Press, 1988.

[51] T. Theodosopoulos. *Worst-Case Identification in ℓ_1: Algorithms and Complexity*, MIT S.M. Thesis, January 1993.

[52] D. Tse, M.A. Dahleh and J.N. Tsitsiklis. Optimal Asymptotic Identification under bounded disturbances. *IEEE Trans. Automat. Contr.*, Vol. 38, No. 8, pp. 1176-1190, August 1993.

[53] D.N.C. Tse, M.A. Dahleh, J.N. Tsitsiklis, " Optimal and Robust Identification in the ℓ_1 norm", in *Proc. of the 1991 American Control Conference*, June 1991.

[54] D.N.C. Tse, " Optimal and robust identification under bounded disturbances", Master's thesis, Dept. of Elec. Eng. and Comp. Sci., M.I.T., February, 1991.

[55] G. Zames, "On the metric complexity of casual linear systems: ϵ-entropy and ϵ-dimension for continuous-time", *IEEE Trans. on Automatic Control*, Vol. 24, April 1979.

[56] Z. Zang and R. Bitmead and M. Gevers, "H_2 Iterative Model Refinement and Control Robustness Enhancement", *Proc. 1991 Conference on Decision and Control*, Brighton, England, December 1991.

Optimal Asymptotic Identification Under Bounded Disturbances

David N. C. Tse, *Student Member, IEEE*, Munther A. Dahleh, *Member, IEEE*, and John N. Tsitsiklis, *Member, IEEE*

***Abstract*—This paper investigates the intrinsic limitation of worst-case identification of LTI systems using data corrupted by bounded disturbances, when the unknown plant is known to belong to a given *model set*. This is done by analyzing the optimal worst-case asymptotic error achievable by performing experiments using *any* bounded inputs and estimating the plant using *any* identification algorithm. First, it is shown that under some topological conditions on the model set, there is an identification algorithm which is asymptotically optimal for any input. Characterization of the optimal asymptotic error as a function of the inputs is also obtained. These results hold for any error metric and disturbance norm. Second, these general results are applied to three specific identification problems: identification of stable systems in the l_1 norm, identification of stable rational systems in the H_∞ norm, and identification of unstable rational systems in the gap metric. For each of these problems, the general characterization of optimal asymptotic error is used to find near-optimal inputs to minimize the error.**

I. Introduction

RECENTLY, there has been a growing line of work with the common theme that system identification should be performed so that the *worst-case* error of the resulting model is small in a metric compatible with robust control [8]–[10], [26], [37]. This paper addresses the questions of asymptotically optimal identification algorithms and experiment designs from this point of view. Our emphasis is less on finding efficient algorithms and more on finding the *fundamental limitations* in identification accuracy achievable by *any* identification algorithm in the limit of observing more and more data corrupted by nonstochastic noise. Thus, this work is in the flavor of the questions posed by Zames [41].

We will deal exclusively with discrete-time, single–input single–output linear time-invariant systems. In this formulation, the unknown plant is *a priori* known to be in a certain subset $\mathfrak{M}$ of the space of all LTI systems; this subset will be called a *model set* $\mathfrak{M}$. The model set is endowed with a general metric ρ which can be any uncertainty measure suitable for designing robust controllers. To identify the plant, one is allowed to perform one or more finite but arbitrarily long experiments using input sequences chosen from a given input set $\mathfrak{U}$. (Typically, $\mathfrak{U}$ is some norm-bounded set.) The measured outputs are corrupted with additive disturbance sequences which are bounded in an l_p norm $\|\cdot\|_p$ but can otherwise be arbitrary. The problem is to analyze the *smallest* worst-case error, over all plants in $\mathfrak{M}$ and all admissible disturbances, achievable by using any inputs from $\mathfrak{U}$ and any identification algorithm to estimate the plant from arbitrarily long but finite data records (i.e., asymptotic error). Our goal is to investigate the key properties of model sets which can be identified with a small optimal error, and in particular how large the model set can be to still yield a finite optimal error. Furthermore, we are interested in robustness issues: does the optimal error vanish as the bound on the output disturbance decreases to zero? Answers to these questions give a characterization of the difficulty of identification using a given model set.

A natural framework to study worst-case identification is provided by information-based complexity theory [21], [35], [36]. This theory provides a general mathematical framework for analyzing the optimal error achievable in solving a problem using a given amount of possibly inaccurate and partial information. *Information* plays the central role in this theory: the results depend only on the information used by an algorithm but are independent of its structure. Our work, like many others in worst-case identification, has employed some of the basic concepts of this theory, but the key results we derived are completely new.

Although mainstream system identification research adopts stochastic models for the noise, there is a line of work which deals with worst-case identification under bounded disturbances [5], [16], [22]–[24], [28], [32], [15]. More recently, specific identification algorithms are proposed in [8]–[10], [26] for worst-case identification in the H_∞ metric from noisy frequency response data and in [12], [25] for identification in the l_1 metric from time series data. In contrast to these works, we deal with general aspects of optimal worst-case asymptotic identification in a general error metric. Moreover, the issue of optimal experiment design, although considered in stochastic system identification (e.g., [7], [20], [43]), has not been satisfactorily addressed in the worst-case setting. Issues of complexity and tradeoffs between the length of experiments and accuracy has been recently reported in [3], [13], [18], [31].

Manuscript received October 23, 1991; October 25, 1992. Paper recommended by Past Associate Editor, T. T. Georgiou. This work was supported in part by an NSERC fellowship from the government of Canada, by the NSF under Grants ECS-8552419, ECS-9157306, and ECS-8552419, by the AFOSR under Grants AFOSR-91-0368 and AFOSR-91-0368, and by the ARO under Grants DAAL03-86-K-0171 and DAAL03-86-K-0171.

The authors are with the Laboratory for Information and Decision Systems, M.I.T., Cambridge, MA 02139.

IEEE Log Number 9210358.

The contributions of this paper are two-folded. At a more general level, it introduces a framework for the analysis of optimal worst-case asymptotic error under bounded disturbances. The central result here is that, under some topological conditions on the model set, *infinite-horizon* experiments, where the entire infinite data record is available to compute estimates, can be viewed as a limit of *finite-horizon* experiments, where only finite data records are available. Analysis of optimal asymptotic error is then reduced to finding optimal inputs to minimize the worst-case error for the infinite-horizon problem. At a more specific level, concrete results are obtained by applying the general framework to three specific identification problems: identification of stable systems in the l_1 and H_∞ metrics, and identification of unstable systems in the gap metric. In all these problems, the required topological conditions for consistency are verified and the infinite-horizon problem is analyzed to find good input designs.

The organization of the paper is as follows. In Section II, the identification problem is formulated and the optimal worst-case asymptotic error achievable by any identification algorithm is defined. In Section III, we present consistency results establishing infinite-horizon experiments as limits of finite-horizon ones. In Section IV, the general results developed are applied to analyze three specific identification problems. Section V contains our conclusion.

II. Problem Formulation

Let $\mathfrak{X}$ be the class of all causal, single-input single-output, linear time-invariant, discrete-time systems. We identify $\mathfrak{X}$ with the space of all one-sided real-valued sequences, R^ω. Let $\mathfrak{M} \subset \mathfrak{X}$ be the *model set* which is assumed to contain the unknown plant h to be identified. The set $\mathfrak{M}$ captures the experimenter's *a priori* knowledge about h. Some examples of $\mathfrak{M}$ are the set of all stable systems, the set of stable systems with a bound on the decay rate, the set of all finite-dimensional systems with a bound on the order, etc. Also given is an input set $\mathfrak{U}$ which contains all the input sequences that can be used in the identification experiments. Typically, $\mathfrak{U}$ is a norm-bounded set, to reflect physical limitations, power restrictions, safety, or to maintain the validity of the linear model of the plant.

An *experiment* is conducted by choosing an input sequence $u \in \mathfrak{U}$ and measuring the output sequence y, related to u by

$$y = h * u + d \tag{2.1}$$

where $*$ denotes the convolution operator and d is the disturbance sequence which corrupts the measurements. (Note that h, u, y, d are all one-sided real-valued sequences: $h = (h_0, h_1, h_2, \cdots$, etc.). The disturbance d is assumed to be bounded in a given norm, $\|d\|_p \le \delta$ for some known δ, but can otherwise be arbitrary. The disturbance may arise from actual measurement noise, such as quantization, or it may reflect nonlinearities and time-variation of the plant. In the latter case, the true plant is actually nonlinear and time varying but is assumed to be approximated well at the operating range by an LTI component, which is the object of identification.

One point to note is that we assume that the system is initially at rest before an experiment is started. Having an unknown nonzero initial condition is equivalent to having an additional, unknown, additive disturbance $u^- * h$, where u^- is the (unknown) input before time $t = 0$. If the model set $\mathfrak{M}$ is bounded in the operator norm from the input space to the disturbance space, then $u^- * h$ is bounded if u^- is, and this additional uncertainty can be accounted for by grouping into the original additive disturbance term. If this is not the case, however, then the problem cannot be treated in the present framework.

Now suppose N such independent experiments are performed. The question whether more than one input is needed to identify plants in a given model set will be addressed. We then have:

$$y^{(i)} = u^{(i)} * h + d^{(i)}, \qquad i = 1, 2, \cdots, N \tag{2.2}$$

where $y^{(i)}$ and $d^{(i)}$ are the output and disturbance sequences in the ith experiment. This can be written in a more compact notation:

$$\boldsymbol{y} = \boldsymbol{u} * h + \boldsymbol{d} \qquad \|\boldsymbol{d}\|_p \equiv \max_i \|d^{(i)}\|_p \le \delta \tag{2.3}$$

where $\boldsymbol{y} = [y^{(1)}, \cdots, y^{(N)}]$, $\boldsymbol{u} = [u^{(1)}, \cdots, u^{(N)}]$, and $\boldsymbol{d} = [d^{(1)}, \cdots, d^{(N)}]$ are vectors of sequences: convolution of h with a vector of inputs is just element-wise convolution with every input. Also note that the vector of inputs $\boldsymbol{u}$ is in $\mathfrak{U}^N$.

An *identification algorithm* is a mapping ϕ which generates, at each time instant n, an estimate $\hat{h}^{(n)} \equiv \phi(P_n \boldsymbol{u}, P_n \boldsymbol{y}) \in \mathfrak{X}$ of the unknown plant h, given the input and output sequences in the experiments. Here, P_n is the truncation operator, defined by $P_n x = (x_0, x_1, \cdots, x_n)$ for each infinite sequence x. Its use signifies that the algorithm ϕ generates at each time instant an estimate based only on the input-output data it has seen so far. Generally, we will assume that the algorithm has access to what the model set $\mathfrak{M}$ is and also the value of δ, the bound on the disturbance. In the terminology of Helmicki *et al.* [12], the algorithm is *tuned*. However, in some cases, we will be able to give stronger results using algorithms which are untuned to the value of δ.

Also given is an extended metric $\rho(\cdot,\cdot)$ on $\mathfrak{X}$, ρ: $\mathfrak{X} \times \mathfrak{X} \to \boldsymbol{R} \cup \{\infty\}$, which evaluates the accuracy of $\hat{h}^{(n)}$ as an estimate of h.

Given an identification algorithm and a chosen set of input sequences for the experiments, we would like to consider the limiting situation when longer and longer of the output sequences are observed. To this end, the worst-case asymptotic error is defined as follows.

Definition 2.1: Fix the inputs $\boldsymbol{u}$. The worst-case asymptotic error, $e_\infty(\phi, \mathfrak{M}, \boldsymbol{u}, \delta)$, of an algorithm ϕ is the smallest number r such that for all plants $h \in \mathfrak{M}$ and for all

disturbances d with $\|d\|_p \le \delta$,

$$\limsup_{n \to \infty} \rho(\phi(P_n u, P_n(u * h + d)), h) \le r.$$

Equivalently,

$$e_\infty(\phi, \mathfrak{M}, u, \delta) \equiv \sup_{h \in \mathfrak{M}} \sup_{\|d\|_\infty \le \delta} \limsup_{n \to \infty} \rho(\phi(P_n u, P_n(u * h + d)), h).$$

According to this definition, no matter what the true plant and the disturbances are, the plant can be eventually approximated to within $e_\infty(\phi, \mathfrak{M}, u, \delta)$, using the estimates generated by the identification algorithm. This is quite analogous to the notion of convergence of estimates to the true plant in the classical probabilistic framework of identification. However, since the disturbances here are assumed to be arbitrary and not necessarily stationary, such convergence is not possible in general. Instead, we only require the estimates to enter and stay within a ball around the true plant rather than to converge to the exact plant.

In the above definition of the worst-case asymptotic error, although convergence of the estimates to within $e_\infty(\phi, \mathfrak{M}, u, \delta)$ is guaranteed for all admissible plants and disturbance sequences, the *rate* of convergence may be arbitrarily slow for some plants and some disturbances. The worst-case asymptotic error is said to be *uniform* if the rate of convergence is uniform over all admissible plants and disturbance sequences. If the convergence is uniform, the worst-case asymptotic error defined above is the same as the limit of the worst-case error taken at each finite time n, i.e.,

$$\sup_{h \in \mathfrak{M}} \sup_{\|d\|_\infty \le \delta} \limsup_{n \to \infty} \rho(\phi(P_n u, P_n(u * h + d)), h) = \limsup_{n \to \infty} \sup_{h \in \mathfrak{M}} \sup_{\|d\|_\infty \le \delta} \rho(\phi(P_n u, P_n(u * h + d)), h)$$

This allows one to *a priori* determine the experiment length required to guarantee that *any* plant in the model set can be identified to a prescribed accuracy. It is the notion of convergence considered by Helmicki *et al.* in their framework [11].

Demanding uniform convergence is too restrictive a formulation for a general theory of fundamental limitations of worst-case identification. Although such uniform convergence is certainly desirable, it is impossible to achieve for many interesting model sets. In fact, for many inherently infinite-dimensional model sets, the worst-case error at each finite time is always infinite, while the worst-case asymptotic error can be made small using an appropriate identification algorithm and inputs. Our formulation thus allows us to discuss optimal worst-case identification and optimal inputs for a much broader class of model sets. Besides, in some applications of identification, such as adaptive control, uniform convergence of estimates is not necessary to fulfill the desired objectives. However, because of the special importance of uniform convergence, we will give additional conditions on the model set for this to take place. It will be seen that these conditions are quite strong and essentially require the model set to be finite-dimensional. It is worthwhile to note that the model set considered in [8], [9] satisfies these conditions.

The optimal worst-case asymptotic error $E_\infty(u, \mathfrak{M}, \delta)$ is defined as the smallest error achievable by any algorithm:

$$E_\infty(u, \mathfrak{M}, \delta) \equiv \inf_\phi e_\infty(\phi, \mathfrak{M}, u, \delta).$$

Any algorithm for which the infimum is attained is said to be *asymptotically optimal*. We will obtain a general characterization of the asymptotically optimal algorithms and the resulting optimal worst-case asymptotic error, for given inputs u. For specific problems, we will find conditions on the inputs u to make this optimal worst-case asymptotic error small.

It should be noted that the asymptotically optimal algorithms to be derived are valid for arbitrary inputs u. This allows the complete separation of the problem of devising optimal algorithms and the problem of designing optimal inputs. This is particularly important when there is no complete control over the choice of the inputs into the plants, such as in closed-loop experiments or in adaptive control. In these problem, this "separation principle" facilitates the derivation of necessary conditions on the input signals for accurate identification to take place.

We would also like to point out that there are some recent asymptotic optimality results in the general information-based complexity framework [14]. However, their notion of optimality is that of the *rate of convergence* of the worst-case error for any *fixed* problem element, and their results only make sense if the error converges to zero. In contrast, in the worst-case identification problem we are dealing with, the error does not typically converge to zero, and our notion of optimality is that of the nonzero limit supremum of the error.

III. Asymptotically Optimal Identification

In this section, the inputs will be assumed to be fixed. The characterization of asymptotically optimal algorithms and optimal worst-case asymptotic error is in terms of the important notion of the *uncertainty set*, an important notion in information-based complexity theory.

Definition 3.1: Let u and y be the input and measured output sequences, and δ be the bound on the disturbances. The finite-horizon uncertainty set at time n is defined to be

$$S_n(\mathfrak{M}, u, y, \delta) = \{g \in \mathfrak{M}: \|P_n(u * g - y)\|_p \le \delta\}$$

and the infinite-horizon uncertainty set is

$$S_\infty(\mathfrak{M}, u, y, \delta) = \{g \in \mathfrak{M}: \|u * g - y\|_p \le \delta\}.$$

The set S_n contains all the plants in the model set consistent with the output data seen until time n. It characterizes the uncertainty at time n: any plant in S_n can be the actual plant from the experimenter's point of view. Similarly, S_∞ contains all the plants that are consis-

tent with the entire output sequences. It measures the uncertainty that the experimenter would still have even if he could perform infinitely long experiments and could see the entire output record. It is easy to see that the finite-horizon uncertainty sets become smaller with increasing n.

For any set $A \subset \mathfrak{X}$, define the diameter and radius of the set A as

$$\operatorname{diam}(A) = \sup_{g,h \in A} \rho(g,h),$$

$$\operatorname{rad}(A) = \inf_{g \in \mathfrak{X}} \sup_{h \in A} \rho(g,h).$$

Note that $\operatorname{diam}(A)/2 \leq \operatorname{rad}(A) \leq \operatorname{diam}(A)$. We shall now define two important quantities.

Definition 3.2: Given a choice of the inputs u, define the infinite-horizon diameter of information $D(u, \mathfrak{M}, \delta)$ and radius of information $R(u, \mathfrak{M}, \delta)$ to be respectively the diameter and radius of the largest possible uncertainty set:

$$D(u, \mathfrak{M}, \delta) \equiv \sup_{h \in \mathfrak{M}} \sup_{\|d\|_\infty \leq \delta} \operatorname{diam}(S_\infty(\mathfrak{M}, u, u * h + d, \delta))$$

$$R(u, \mathfrak{M}, \delta) \equiv \sup_{h \in \mathfrak{M}} \sup_{\|d\|_\infty \leq \delta} \operatorname{rad}(S_\infty(\mathfrak{M}, u, u * h + d, \delta)).$$

In information-based complexity terminology, these quantities correspond to the *diameter* and *radius of information* for the infinite-horizon problem where the information available is the entire infinite output sequence. The quantity $D(u, \mathfrak{M}, \delta)$ is the largest distance between two plants for which there are admissible disturbances such that the plants give exactly the same outputs. It turns out that it is precisely this quantity that characterizes the optimal worst-case asymptotic errors. First we show that half the infinite-horizon diameter of information is a lower bound to the optimal asymptotic error.

Proposition 3.3: Let $\mathfrak{M}$ be any model set, u be any vector of inputs and $\delta \geq 0$. Then

$$e_\infty(\phi, \mathfrak{M}, u, \delta) \geq D(u, \mathfrak{M}, \delta)/2$$

for any algorithm ϕ.

Proof: Let ψ be an algorithm for the infinite-horizon problem, i.e., given the entire input and output sequences, ψ generates an estimate for the plant. The worst-case error achieved by this algorithm is:

$$\sup_{h \in \mathfrak{M}} \sup_{\|d\|_\infty \leq \delta} \rho(\psi(u, u * h + d), h)$$

and the infinite-horizon optimal worst-case error achievable by any algorithm is

$$\inf_{\psi} \sup_{h \in \mathfrak{M}} \sup_{\|d\|_\infty \leq \delta} \rho(\psi(u, u * h + d), h). \tag{3.4}$$

One should note that while the algorithms allowed in this infinite-horizon problem have access to the entire infinite input–output sequences, the algorithms for the asymptotic problem have access to only finite but arbitrarily long portions. Consequently, the infinite-horizon optimal worst-case error lower bounds the optimal asymptotic error $E_\infty(u, \mathfrak{M}, \delta)$. On the other hand, by a central result in information-based complexity theory [35], this infinite-horizon optimal error is given by the infinite-horizon radius of information $R(u, \mathfrak{M}, \delta)$, which in turn is lower bounded by half the diameter of information $D(u, \mathfrak{M}, \delta)$. Hence, the result follows. □

The key question now is whether there exists an optimal algorithm which can always generate estimates with error converging to this lower bound. By the definition of the infinite-horizon uncertainty set, there exist two plants at a separation of $D(u, \mathfrak{M}, \delta)$ which can give rise to exactly the same output measurements. Thus in the worst case, there is no way for any finite-duration experiments to distinguish between them, and this gives rise to the lower bound proved above. Conversely, any two plants with a separation greater than $D(u, \mathfrak{M}, \delta)$ can be distinguished if we perform experiments of sufficiently long length. That is, if h is the true plant, and h' is another plant which is far away from h (separation greater that $D(u, \mathfrak{M}, \delta)$), there exists a time $T(h')$ for which one needs to observe the output to eliminate h' from consideration as a possible candidate. However, to guarantee that an accurate estimate at time n can be obtained, one needs $T(h') < n$ for *all* plants h' that are far away from h. Otherwise, although the identification algorithm always picks estimates which are consistent with the output seen so far, the estimates may nevertheless diverge from the true plant.

The issue discussed above is really one of *consistency* between finite-horizon experiments, where only a finite data record is available for computing estimates, and infinite-horizon experiments, where the entire infinite data record is available. The question is when the latter can be viewed as a limit of the former. In [17], such a consistency result is established by placing a stationarity assumption on the noise and then appealing to the law of large numbers. As far as we know, this issue has not been considered in an unknown-but-bounded noise setting. In fact, it will now be shown that a compactness condition on the model set will guarantee consistency.

The following theorem shows that, under a σ-compactness assumption on $\mathfrak{M}$, $D(u, \mathfrak{M}, \delta)$ is an upper bound for the optimal asymptotic error. Combining with Proposition 3.3, we have upper and lower bounds that agree, within a factor of 2. Thus, the study of the optimal asymptotic error is reduced to the study of $D(u, \mathfrak{M}, \delta)$, if we ignore this factor of 2.

Theorem 3.4: Suppose that the model set $\mathfrak{M}$ is σ-compact in the ρ-topology, $\mathfrak{M} = \bigcup_i \mathfrak{M}_i$, $M_i \subset M_{i+1}$ $\forall_i$, $\mathfrak{M}_i$ compact and on each $\mathfrak{M}_i$, convergence in the ρ-topology implies component-wise convergence of the impulse response. Then there is an identification algorithm ϕ^* such that $e_\infty(\phi^*, \mathfrak{M}, u, \delta) \leq D(u, \mathfrak{M}, \delta)$ for all u and $\delta \geq 0$.

It should be noted that by an elementary result in information-based complexity theory, the optimal worst-case error achievable when the algorithm has *full* access to the entire *infinite* input–output sequences is also

bounded between the infinite-horizon diameter of information and half the diameter of information. Our two results (Proposition 3.3 and Theorem 3.4) are of an entirely different nature: they assert that the optimal worst-case asymptotic error achievable when the algorithm has access to *finite* but arbitrarily long data records also satisfies the same bounds. The assumed topological conditions are crucial for the validity of Theorem 3.4.

Before proving Theorem 3.4, we need one more definition and a few lemmas.

Definition 3.5: For given inputs u and bound δ on disturbances, and $g, h \in \mathfrak{X}$, define $T_{u,\delta}(g,h)$ to be the smallest integer k such that $\|P_k(u*(g-h))\|_p > 2\delta$. If no such k exists, then $T_{u,\delta}(g,h)$ is infinite.

Lemma 3.6: For any two plants $g, h \in \mathfrak{M}$, $T_{u,\delta}(g,h)$ is the smallest k such that there is no output y with g and h in the same uncertainty set $S_k(\mathfrak{M}, u, y, \delta)$.

Proof: If $n = T_{u,\delta}(g,h)$, then $\|P_n\{u*(g-h)\}\|_p > 2\delta$, so for every output sequence y, either $\|P_n\{u*g - y\}\|_p > \delta$ or $\|P_n\{u*h - y\}\|_p > \delta$, by the triangle inequality. Hence, g and h cannot be in the same uncertainty set $S_n(\mathfrak{M}, u, y, \delta)$ for any y. Conversely, if $n < T_{u,\delta}(g,h)$, then $P_n\{u*(g-h)\}\|_p \le 2\delta$, so picking $y = u*(g-h)/2$ yields $\|P_n\{u*g - y\}\|_p \le \delta$ and $\|P_n\{u*h - y\}\|_p \le \delta$. Hence, $g, h \in S_n(\mathfrak{M}, u, y, \delta)$. □

Thus, given two plants g and h, $T_{u,\delta}(g,h)$ is the minimum duration for which one has to observe the output to ensure that at least one of the two plants can be eliminated from consideration as the true plant.

Lemma 3.7: Let $g, h \in \mathfrak{M}$. If $\rho(g,h) > D(u, \mathfrak{M}, \delta)$, then $T_{u,\delta}(g,h) < \infty$.

Proof: Suppose $T_{u,\delta}(g,h) = \infty$. Then $\|P_k\{u*(g-h)\}\|_p \le 2\delta$ for every k, so $\|u*(g-h)\|_p \le 2\delta$. Now consider the disturbance $d = u*(h-g)/2$, and the infinite-horizon uncertainty set $S_\infty(\mathfrak{M}, u, u*g+d, \delta)$, arising when g is the true plant. (Note that $\|d\|_p \le \delta$.) But $\|u*h - (u*g+d)\|_p = \|u*(h-g)/2\|_p \le \delta$, so the plant h is also in the set $S_\infty(\mathfrak{M}, u, u*g+d, \delta)$. Hence, by definition of the infinite-horizon diameter of information, $\rho(g,h) \le D(u, \mathfrak{M}, \delta)$. □

The desired topological condition involves the topology of component-wise convergence of sequences, or the so-called product topology [27].

Lemma 3.8: Fix the inputs $u \in \overline{Bl}_\infty^N$ and $\delta > 0$. Let $A \subset \mathfrak{M} \times \mathfrak{M}$ be compact in the product topology, and suppose $T_{u,\delta}(g,h)$ is finite for every $(g,h) \in A$. Then $\sup_{(g,h)\in A} T_{u,\delta}(g,h)$ is also finite.

Proof: Suppose $\sup_{(g,h)\in A} T_{u,\delta}(g,h) = \infty$. Then there exists a sequence of plants $(g^{(i)}, h^{(i)})$ in A such that $\lim_{i\to\infty} T_{u,\delta}(g^{(i)}, h^{(i)}) = \infty$; furthermore, the sequence can be assumed to converge (in the product topology) to a pair of plant $(g^*, h^*) \in A$ since A is compact. Let $n^* \equiv T_{u,\delta}(g^*, h^*) < \infty$. By definition, $\|P_{n^*}(u*(g^*-h^*))\|_p > 2\delta$. Since the norm of a sequence is a continuous function of finitely many of its components, it follows that $\|P_{n^*}(u*(g-h))\|_p$ is a continuous function of (g,h) in the product topology. Hence, there exists a ball B (in the product topology) around (g^*, h^*) such that for every $(g', h') \in B$, $\|P_{n^*}(u*(g'-h'))\|_p > 2\delta$, i.e., $T_{u,\delta}(g',h') \le n^*$ for every $(g',h') \in B$. But this contradicts the fact that $\lim_{i\to\infty} T_{u,\delta}(g^{(i)}, h^{(i)}) = \infty$ since $(g^{(i)}, h^{(i)}) \to (g^*, h^*)$. Hence, it can be concluded that $\sup_{(g,h)\in A} T_{u,\delta}(g,h)$ is in fact finite. □

Basically, this lemma says that if each plant in the compact set A can be eventually ruled out as the true plant, there is a finite time after which all of them can be simultaneously ruled out.

Now we are in a position to prove Theorem 3.4.

Proof: Define the identification algorithm ϕ^* as follows: at each time n, the algorithm generates as an estimate by picking any arbitrary plant $\hat{h}^{(n)}$ in the set $S_n \cap \mathfrak{M}_k$, where S_n is the uncertainty set after observing the output data until time n, and k is the least integer i such that $S_n \cap \mathfrak{M}_i$ is nonempty. We claim that this algorithm will have an asymptotic error of at most $D(u, M, \delta)$ for all inputs u and $\delta > 0$.

Fix the unknown plant $h \in \mathfrak{M}$ and let $\epsilon > 0$. Also let $\mathfrak{M}_h$ be the smallest of the compact subsets $\mathfrak{M}_i$'s which contains h. Define the set

$$A(h, \epsilon) \equiv \{g \in \mathfrak{M}_h : \rho(g,h) \ge D(u, \mathfrak{M}, \delta) + \epsilon\} \quad (3.5)$$

and the number

$$T(h, \epsilon) \equiv \sup_{g \in A(h,\epsilon)} T_{u,\delta}(g,h). \quad (3.6)$$

Since $A(h, \epsilon)$ is a closed subset of $\mathfrak{M}_h$ (with respect to the ρ-topology), it is also compact in the ρ-topology. Since the ρ-topology is finer than the product topology in $\mathfrak{M}_h$, $A(h, \epsilon)$ is also compact in the product topology. By Lemma 3.7, $T_{u,\delta}(g,h)$ is finite for all $(g,h) \in A(h,\epsilon)$. Hence, by Lemma 3.8, $T(h, \epsilon)$ is also finite.

Now consider the estimates $\hat{h}^{(n)}$ generated by the algorithm ϕ^*. Since $\hat{h}^{(n)}$ is picked from the least k such that $S_n \cap \mathfrak{M}_k$ is nonempty, $\hat{h}^{(n)}$ is guaranteed to be in M_h for all n. (This is because $S_n \cap \mathfrak{M}_h$ is nonempty: it contains the true plant h.) Also $\hat{h}^{(n)}$ is in the uncertainty set S_n and by Lemma 3.6, $T_{u,\delta}(\hat{h}^{(n)}, h) > n$. If we now take any $n > T(h, \epsilon)$, we have $T_{u,\delta}(\hat{h}^{(n)}, h) > T(h,\epsilon)$ so $\hat{h}^{(n)}$ is not in $A(h, \epsilon)$. But $\hat{h}^{(n)}$ is in $\mathfrak{M}_h$, so it follows that $\rho(\hat{h}^{(n)}, h) < D(u, \mathfrak{M}, \delta) + \epsilon$.

Since ϵ is arbitrary, it can now be concluded that

$$\limsup_{n\to\infty} \rho(h^{(n)}, h) \le D(u, \mathfrak{M}, \delta)$$

completing the proof. □

The above construction of the asymptotically near-optimal algorithm ϕ^* can be viewed as an application of *Occam's Razor*—that one should always use the "simplest" theory to explain the given data. Here, as is true in general, there is no absolute measure of simplicity. Rather it is defined by the choice of the nested partitioning of the model set, $\mathfrak{M} = \bigcup_i \mathfrak{M}_i$. Given this nested structure, plants in the smaller $\mathfrak{M}_i$'s are considered to be simpler than those in larger $\mathfrak{M}_i$. Convergence of the estimates is guaranteed by always choosing the simplest plant that is consistent with the data seen so far. This avoids overfitting of

data, a problem which crops up all the time in statistics and pattern recognition. It is interesting to note that this same principle of Occam's Razor has also been applied to guarantee convergence in distribution-free probabilistic learning problems [1], [30].

In contrast to the σ-compactness condition that guarantees convergence, a stronger compactness condition guarantees uniform convergence.

Proposition 3.9: Suppose convergence in the ρ-topology on $\mathfrak{M}$ implies component-wise convergence of the impulse response. If the model set $\mathfrak{M}$ is compact in the ρ-topology, then there is an algorithm ϕ the estimates of which will converge uniformly to within $D(u, \mathfrak{M}, \delta)$ of the true plant: i.e., for all $\epsilon > 0$, there exists a time $T(\epsilon)$ such that for all $h \in \mathfrak{M}$, $\|d\|_p \leq \delta$,

$$\rho(\phi(P_n u, P_n(u * h + d)), h) \leq D(u, \mathfrak{M}, \delta) + \epsilon \qquad \forall n > T(\epsilon).$$

Moreover, the algorithm does not require the knowledge of δ, the bound on the disturbances, to compute its estimates.

Proof: An algorithm ϕ is defined as follows: for each n,

$$\phi(P_n u, P_n y) = \underset{g \in \mathfrak{M}}{\operatorname{argmin}} \|P_n(u * g - y)\|_p. \qquad (3.7)$$

The minimum must exist since $\mathfrak{M}$ is compact and $\|P_n(u * g - y)\|_p$ is a continuous function of g in the product topology and hence in the ρ-topology. Also note that computing this estimate does not require the knowledge of δ.

Now $y = u * h + d$ for some true plant h and disturbance d satisfying $\|d\|_p \leq \delta$. By definition, the estimate at each time n satisfies

$$\|P_n(u * \phi(P_n u, P_n y) - y)\|_p \leq \|P_n(u * h - y)\|_p = \|P_n d\|_p \leq \delta$$

and hence $\phi(P_n u, P_n y) \in S_n(\mathfrak{M}, u, y, \delta)$ for each n, where S_n is the finite-horizon uncertainty set at time n. We shall use only this property of the estimates of ϕ to show that they uniformly converge.

Let $\epsilon > 0$. For each plant $h \in \mathfrak{M}$, define

$$A(h, \epsilon) \equiv \{g \in \mathfrak{M}: \rho(g, h) \geq D(u, \mathfrak{M}, \delta) + \epsilon\}. \quad (3.8)$$

Also, consider the number

$$T(\epsilon) \equiv \sup_{h \in \mathfrak{M}} \sup_{g \in A(h, \epsilon)} T_{u, \delta}(g, h) \qquad (3.9)$$

where the function $T_{u, \delta}$ has been defined earlier. $T(\epsilon)$ can be rewritten as $\sup_{(g, h) \in B(\epsilon)} T_{u, \delta}(h')$, where

$$B(\epsilon) = \{(g, h) \in \mathfrak{M}^2: \rho(g, h) \geq D(u, \mathfrak{M}, \delta) + \epsilon\}.$$

It is clear that $B(\epsilon)$ is a closed set and hence compact in the ρ-topology, being a subset of $\mathfrak{M}^2$. Hence, $B(\epsilon)$ is also compact in the product topology. Now, $T_{u, \delta}(g, h)$ is finite for all (g, h) in $B(\epsilon)$, by Lemma 3.7. Hence, by Lemma 3.8, $T(\epsilon)$ is finite.

Now if $n > T(\epsilon)$, then for any plant $h \in \mathfrak{M}$ and $\|d\|_p \leq \delta$, the estimate $\hat{h}^{(n)}$ generated by the algorithm must lie in the uncertainty set $S_n(\mathfrak{M}, u, u * h + d, \delta)$. Hence, by Lemma 3.6, $T_{u, \delta}(\hat{h}^{(n)}, h) > n > T(\epsilon)$. This implies

$$\rho(\hat{h}^{(n)}, h) < D(u, \mathfrak{M}, \delta) + \epsilon.$$

Since this holds for all h and d, the convergence is indeed uniform. □

IV. Application of General Framework to Specific Problems

The above results state that under some compactness conditions on the model set, the optimal worst-case asymptotic error achievable by any identification algorithm is characterized by the function $D(u, \mathfrak{M}, \delta)$, measuring the worst-case uncertainty from infinite-horizon experiments. It describes the intrinsic difficulty of identifying plants in a given model set, independent of the specific identification algorithm used. This result enables us to move from the analysis of the error of specific algorithms to the analysis of the function $D(u, \mathfrak{M}, \delta)$. In specific problems, we would like to find inputs u such that $D(u, \mathfrak{M}, \delta)$ is small or, at the very least, vary continuously with the noise bound δ at $\delta = 0$. This would imply that identification accuracy is *robust* to measurement noise.

The value of the diameter of information $D(u, \mathfrak{M}, \delta)$ is in general difficult to evaluate because it is the supremum over the diameter of all possible infinite-horizon uncertainty sets. However, if the ρ metric comes from a norm, it turns out that for an important class of model sets, $D(u, \mathfrak{M}, \delta)$ has a simple characterization. These are the model sets which are convex and balanced. (A set A is said to be balanced if for every h in A, $-h$ is also in A.) The following proposition gives the characterization, and it follows from a basic result in information-based complexity theory [21].

Proposition 4.1: Suppose $\rho(g, h) \equiv \|g - h\|_{\mathfrak{X}}$ for some norm $\|\cdot\|_{\mathfrak{X}}$. If $\mathfrak{M}$ is a balanced convex subset of $\mathfrak{X}$, then the worst-case diameter is attained when the true plant and the disturbance are both 0. That is,

$$D(u, \mathfrak{M}, \delta) \equiv \sup_{h \in \mathfrak{M}} \sup_{\|d\|_\infty \leq \delta} \operatorname{diam}(S_\infty(\mathfrak{M}, u, u * h + d, \delta)) = \operatorname{diam}(S_\infty(\mathfrak{M}, u, 0, \delta)).$$

Now we will apply the general results proved above to analyze specific identification problems. We take our input set $\mathfrak{U}$ to be $\overline{Bl}_\infty \equiv \{u: \|u\|_\infty \leq 1\}$, where $\|u\|_\infty \equiv \sup_i |u_i|$. (The 1 is taken for normalization purpose.) The disturbance is assumed to be an l_∞ signal d, with $\|d\|_\infty \leq \delta$.

A. Identification of Stable Plants in the l_1 Norm

Here the metric considered is $\rho(g, h) \equiv \|g - h\|_1$, and we restrict ourselves to stable plants with impulse responses of finite l_1 norm. We shall first prove a general lower bound for $D(u, \mathfrak{M}, \delta)$ which holds for all inputs u and for a wide class of model sets.

Proposition 4.2: Assume the model set $\mathfrak{M}$ contains two plants at an l_1 distance of 2δ apart. Then for any number of experiments N and any set of inputs $u \in \overline{Bl}_\infty^N$,

$$D(u, \mathfrak{M}, \delta) \geq 2\delta.$$

Proof: Let $g, h \in \mathfrak{M}$ satisfy $\|g - h\|_1 = 2\delta$. Suppose that u are the inputs used in the identification experiments and h is the actual plant. Let the disturbance be $d = u*(g - h)/2$. Note that $\|d\|_\infty \leq \|u\|_\infty \|(g - h)/2\|_1 = \delta$.

The observed output is $y = u*h + d = u*(g + h)/2$. Now, $\|u*g - y\|_\infty = \|(1/2)u*(g - h)\|_\infty \leq (1/2)\|u\|_\infty \|g - h\|_1 \leq \delta$. Therefore, $g \in S_\infty(\mathfrak{M}, u, y, \delta)$. Since h is also in $S_\infty(\mathfrak{M}, u, y, \delta)$, it follows that

$$\text{diam}\,(S_\infty(\mathfrak{M}, u, y, \delta)) \geq \|g - h\|_1 = 2\delta.$$

Since $D(u, \mathfrak{M}, \delta)$ is the diameter of the largest possible uncertainty set, the desired lower bound follows. $\square$

We now demonstrate that in fact, for *all* balanced and convex model sets of stable plants, this lower bound can be reached using just *one* input, provided that it satisfies a persistent excitation property.

Definition 4.3: Let $\mathfrak{A}$ be the set of all finite sequences of 1's and -1's:

$$\mathfrak{A} \equiv \{(a_1, a_2, \cdots, a_k): k \geq 1, a_i \in \{1, -1\}, \forall i\}. \quad (4.10)$$

The sequence $v \in \overline{Bl}_\infty$ is said to contain all finite sequences of 1's and -1's if for every finite sequence $a \in \mathfrak{A}$, there exist m, n such that $(v_m, v_{m+1}, \cdots, v_{m+n}) = a$.

Theorem 4.4: Assume $\mathfrak{M}$ is balanced and convex and contains only stable plants. If u^* contains all finite sequences of 1's and -1's, then

$$D(u^*, \mathfrak{M}, \delta) \leq 2\delta.$$

Proof: By Proposition 4.1, the diameter of information is given by the diameter of the uncertainty set centered at 0:

$$D(u^*, \mathfrak{M}, \delta) = \text{diam}\,(S_\infty(\mathfrak{M}, u^*, 0, \delta)).$$

Consider any $g \in S_\infty(\mathfrak{M}, u^*, 0, \delta)$ and let $\epsilon > 0$. Since g is stable, there exists M such that

$$\sum_{k=M+1}^{\infty} |g_k| < \epsilon. \quad (4.11)$$

Now consider the finite sequence

$$(\text{sgn}\,(g_M), \text{sgn}\,(g_{M-1}), \cdots, \text{sgn}\,(g_0)) \in \mathfrak{A}$$

where sgn is the signum function such that $\text{sgn}\,(x) = 1$ if $x \geq 0$ and $\text{sgn}\,(x) = -1$ if $x < 0$.

By definition of the sequence u^*, there exists m such that

$$u_m^* = \text{sgn}\,(g_M), u_{m+1}^* = \text{sgn}\,(g_{M-1}), \cdots,$$
$$u_{m+M}^* = \text{sgn}\,(g_0).$$

We then have

$$\begin{aligned} |(u^* * g)_{m+M}| &= \left| \sum_{k=0}^{m+M} u^*_{m+M-k} g_k \right| \\ &= \left| \sum_{k=0}^{M} u^*_{m+M-k} g_k + \sum_{k=M+1}^{m+M} u^*_{m+M-k} g_k \right| \\ &= \left| \sum_{k=0}^{M} \text{sgn}\,(g_k) g_k + \sum_{k=M+1}^{m+M} u^*_{m+M-k} g_k \right| \\ &\geq \sum_{k=0}^{M} |g_k| - \sum_{k=M+1}^{m+M} |g_k| \\ &\geq \|g\|_1 - \epsilon. \end{aligned} \quad (4.12)$$

But $g \in S_\infty(\mathfrak{M}, u^*, 0, \delta)$, so $|(u^* * g)_{m+M}| \leq \delta$. Hence, it follows from inequality (4.12) that $\|g\|_1 \leq \delta + \epsilon$. Since this is true for every $\epsilon > 0$, it follows that $\|g\|_1 \leq \delta$ for any $g \in S_\infty(\mathfrak{M}, u^*, 0, \delta)$. Thus,

$$\begin{aligned} D(u, \mathfrak{M}, \delta) &= \text{diam}\,(S_\infty(\mathfrak{M}, u^*, 0, \delta)) \\ &= \sup_{g \in S_\infty(\mathfrak{M}, u^*, 0, \delta)} 2\|g\|_1 \leq 2\delta. \end{aligned} \quad \square$$

An input satisfying the above condition has been proposed independently by Makila [25] for l_1 identification. It is also of interest to note that the random binary sequence, a commonly used identification input generated by randomly and independently picking each value to be 1 or -1, has the desired property of containing all finite sequences of 1's and -1's, with probability 1.

Using the above result on the infinite-horizon diameter of information, we shall analyze the optimal asymptotic l_1 error for stable model sets.

The consistency result proved earlier applies to σ-compact model sets. The following technical lemma concerning the asymptotic l_1 error enables us to extend the result to model sets which are closure of σ-compact model sets as well.

Lemma 4.5: For any model set $\mathfrak{M}$, inputs $u \in \overline{Bl}_\infty^N$, algorithm ϕ and $\delta \geq 0$,

$$e_\infty^1(\phi, u, \overline{\mathfrak{M}}, \delta) \leq \lim_{x \downarrow \delta} e_\infty^1(\phi, u, \mathfrak{M}, x)$$

where $\overline{\mathfrak{M}}$ is the closure of $\mathfrak{M}$ with respect to the l_1-topology on $\mathfrak{X}$. (The superscript "1" emphasizes that the metric used is the l_1 norm.)

Proof: By definition, for all $x \geq 0$, and $\forall h \in \mathfrak{M}$ and d with $\|d\|_\infty \leq x$, we have

$$\limsup_{n \to \infty} \|\phi(P_n u, P_n(u*h + d)) - h\|_1 \leq e_\infty^1(\phi, u, \mathfrak{M}, x). \quad (4.13)$$

Let $\epsilon > 0$. Take any $h \in \mathfrak{M}$ and $\|d\|_\infty \leq \delta$. There exists a $h' \in \mathfrak{M}$ such that $\|h - h'\|_1 \leq \epsilon$. Therefore

$$\limsup_{n\to\infty} \|\phi(P_n u, P_n(u * h + d)) - h\|_1 \tag{4.14}$$

$$\leq \limsup_{n\to\infty} \|\phi(P_n u, P_n(u * h' + u*(h - h') + d)) - h'\|_1 + \epsilon. \tag{4.15}$$

Now, $\|u*(h - h') + d\|_\infty \leq \delta + \epsilon$, so applying inequality (4.13) with $x = \delta + \epsilon$,

$$\limsup_{n\to\infty} \|\phi(P_n u, P_n(u * h' + u*(h - h') + d)) - h'\|_1 \leq e_\infty^1(\phi, \mathfrak{M}, u, \delta + \epsilon). \tag{4.16}$$

It follows that

$$\limsup_{n\to\infty} \|\phi(P_n u, P_n(u * h + d)) - h\|_1 \leq e_\infty^1(\phi, \mathfrak{M}, u, \delta + \epsilon) + \epsilon. \tag{4.17}$$

Letting ϵ go to 0 gives the desired result. □

We now show that we can get very good asymptotic error even if there is no additional prior knowledge about the plant other than the fact that it is stable.

Proposition 4.6: Take the model set to be l_1, the space of all stable plants. There is a single experiment, using any input $u^* \in Bl_\infty$ containing all sequences of 1's and -1's, such that for every $\delta \geq 0$, the optimal asymptotic l_1 error satisfies

$$E_\infty^1(u^*, l_1, \delta) \leq 2\delta.$$

Proof: The space l_1 is separable, i.e., it is a closure of a countable set $\mathfrak{M}_\infty$. Since a countable set is clearly σ-compact, by Theorem 3.4, there is an algorithm ϕ^* such that for every $\delta \geq 0$ and inputs u,

$$e_\infty^1(\phi^*, \mathfrak{M}_\infty, u, \delta) \leq D(u, \mathfrak{M}_\infty, \delta). \tag{4.18}$$

Now, using any input u^* containing all sequences of 1's and -1's, we have

$$\begin{aligned} &e_\infty^1(\phi^*, \mathfrak{M}_{\text{stab}}, u^*, \delta) \\ &\leq \lim_{x\downarrow\delta} e_\infty^1(\phi^*, \mathfrak{M}_\infty, u^*, x) \qquad \text{by Proposition 4.5} \\ &\leq \lim_{x\downarrow\delta} D(u^*, \mathfrak{M}_\infty, x) \\ &\leq 2\delta. \qquad \text{by Theorem 4.4.} \end{aligned}$$ □

Hence, to identify a plant accurately in the limit, it is enough to know *a priori* that it is stable; no additional information, such as bounds on decay rate and gain, is necessary. The achievable accuracy varies continuously with the noise bound δ for small δ; thus, identification can be performed robust to measurement noise. One should also note that there are many other choices of decomposing the model set into compact sets. The decomposition should be done to facilitate a more efficient implementation of the identification algorithm. We will discuss this at the end of this section.

Next, we look at the issue of uniform convergence. For the model set l_1, it can at once be seen that although convergence to a small asymptotic error is possible, such convergence cannot be uniform.

Proposition 4.7: Let ϕ be any algorithm and u be any input. Then for every n and for every M, there exists an $h \in \mathfrak{M}_{stab}$ such that

$$\|\phi(P_n u, P_n(u * h)) - h\|_1 > M.$$

Proof: This is clear because making n measurements gives no information on the part of the impulse response after time n, which can have arbitrarily large uncertainty in the l_1 norm. □

To guarantee uniform convergence, we need to look at compact model sets.

Proposition 4.8: Let $\mathfrak{M} \subset \mathfrak{M}_{\text{stab}}$ be a compact set (in the l_1-topology) or a subset of a compact set in M_{stab}. For the single input u^* which contains all finite sequences of 1's and -1's, there is an algorithm the estimates of which converge, uniformly for all $h \in \mathfrak{M}$ and all $\|d\|_\infty \leq \delta$, to an l_1 ball of radius 2δ around the true plant. Moreover, the algorithm does not require the knowledge of the value of δ to compute its estimates.

Common examples of such compact model sets are the uniformly stable ones, of the form $M_s(g) \equiv \{h: |h_i| \leq |g_i| \text{ for all } i\}$ where g is any stable plant. The specific model sets considered in [8] and [9] belong to this class.

Identification Algorithms for Stable Plants: For certain parameterizations of the space of stable plants, it is possible to device algorithms based on the *Occam's Razor Principle* that involve linear programming problems. Define the compact sets:

$$\mathfrak{M}_k = \{h \in l_1 : |h_i| \leq kM, h_i = 0 \; \forall i \geq k\}$$

and M is any positive real number. It can be immediately seen that

$$l_1 = \text{closure of } \bigcup_{k=1}^{\infty} \mathfrak{M}_k.$$

Fix some tolerance level ϵ. The estimator can be described as picking a feasible element in the set

$$\mathfrak{M}_k \cap S_n(\mathfrak{M}, u, y, \delta + \epsilon)$$

for any input–output pair. Of course, this set is characterized by linear constraints and finding a feasible plant is equivalent to solving a linear programming problem. The estimate is picked from the smallest $\mathfrak{M}_k$ for which the above set is not empty.

Suppose that the model set is equal to $\mathfrak{M}_s(g)$ where $g \in l_1$ and $g_i = 0 \;\; \forall i \geq l$. This set contains only FIR plants of length l, with a bound on the impulse response. For this model set the near-optimal algorithm ϕ^* is given by

$$\phi^*(P_n u, P_n y) = \operatorname*{arg\,min}_{|h_i| \leq |g_i|,\, i=0,1,\cdots l} \|P_n(y - u * h)\|_\infty$$

which is computable by linear programming. We finally note that work on algorithms is still an active area of research [34].

B. H_∞ Identification of Stable Rational Plants

We now analyze optimal identification using the model set RH_∞, space of all stable plants with rational transfer functions. The error metric used is the H_∞ norm. The model set RH_∞ is σ-compact in the H_∞-topology. (For example, it can be decomposed as a countable union of compact sets of the form $\{h: |h_n| \leq A\alpha^n\}$ with A tending to infinity and α tending to 1.) Convergence in H_∞ implies component-wise convergence of the impulse response in each of these sets. Hence, the consistency result applies and we are reduced to the analysis of the infinite-horizon diameter of information.

Since the H_∞-norm of a plant is always upper bounded by its l_1 norm, Theorem 4.4 implies that, measured in the H_∞ norm, the infinite horizon diameter of information $D(u^*, RH_\infty, \delta)$ using an input u^* containing all finite sequences of 1's and -1's is also bounded by 2δ. Hence, the worst-case asymptotic error using this input is also bounded by 2δ. The following result shows that this input is optimal to within a factor of two.

Proposition 4.9: For any number of experiments N and any choice of inputs $\boldsymbol{u} \in \overline{Bl}_\infty^N$, the H_∞ infinite-horizon diameter of information satisfies:

$$D(\boldsymbol{u}, RH_\infty, \delta) \geq 2\delta.$$

Proof: The proof is trivial. Take $g = (\delta, 0, 0, \cdots)$, $h = (-\delta, 0, 0, \cdots)$, $\boldsymbol{d} = -\delta\boldsymbol{u}$, $\boldsymbol{d}' = \delta\boldsymbol{u}$. Then $\boldsymbol{u} * g + \boldsymbol{d} = \boldsymbol{u} * h + \boldsymbol{d}'$ so $D(\boldsymbol{u}, RH_\infty, \delta) \geq \|g - h\|_{H_\infty} = 2\delta$. □

A similar result on frequency response experiments is given by [9].

C. Identification of Unstable Plants in the Gap Metric

Our general framework of optimal asymptotic identification applies, to a large extent, to unstable as well as stable systems. In particular, the consistency and uniform convergence results, for arbitrary inputs, hold regardless of whether the model set contains stable or unstable systems. There is, however, an important issue in the identification of unstable systems which is not dealt with in this framework. While stable systems can be identified in the open-loop, identification experiments for unstable systems are almost always performed in the closed-loop to avoid unbounded outputs. As opposed to open-loop identification, there is no complete freedom in choosing the inputs u for closed-loop identification experiments, as there is a coupling between the input and the output. This makes the experiment design problem much more difficult. In this section, we shall ignore the coupling and confine ourselves to deriving necessary and sufficient conditions on the inputs for accurate asymptotic identification of unstable systems. The question of whether one can design closed-loop experiments to achieve such conditions is left open.

An appropriate error metric to use for unstable plants is the *gap metric* [6], [33], [42]. The important property of the gap metric is that it generates the graph topology [40], which is the weakest topology in which closed-loop stability is a robust property, or in which the closed-loop system varies continuously as a function of the open-loop system. Intuitively, this means that identifying plants accurately in the gap metric is the least that one must do to be able to design controllers to guarantee that the closed-loop performance will be close to the desired.

The gap between two possibly unstable plants is given in terms of their *graphs*, so we will first define this notion. The graph G_h of a plant h is a subset of the space $l_2 \times l_2$, defined by

$$G_h \equiv \{(x, h * x): x \in l_2, h * x \in l_2\}.$$

Thus, the graph of a plant describes its behavior on bounded-energy inputs which yield bounded-energy outputs. The directed gap between two graphs G_h and G_g is defined as

$$\vec{\delta}(G_h, G_g) \equiv \sup_{x \in G_h, \|x\|_2 \leq 1} \inf_{y \in G_g} \|x - y\|_2.$$

The gap between two plants is given by the maximum of the two directed gaps between the two graphs:

$$\delta(g, h) \equiv \max\left(\vec{\delta}(G_g, G_h), \vec{\delta}(G_h, G_g)\right).$$

It can be verified that the gap is indeed a metric, and that its value is always bounded between 0 and 1.

In the analysis below, we shall restrict ourselves to the space of finite-dimensional systems, $\mathfrak{M}_{fd}$, with rational z-transform.[1] In this space, convergence in the graph topology can be expressed in terms of the coprime factors: $P_i \to P$ in the graph topology iff there exist co-prime factorizations $P_i = N_i/D_i$, $P = N/D$ such that $N_i \to N$ and $D_i \to D$ in the H_∞-topology. Results obtained for finite-dimensional plants are also valid for infinite-dimensional systems that can be approximated by finite-dimensional systems in the gap metric.

To apply the consistency results we proved earlier, we have to investigate the topological properties of M_{fd}.

Proposition 4.10: Let p, q be nonnegative integers, k, α be positive real numbers and $\mathfrak{M}_{fd}(p, q, K, \alpha)$ be the class of all finite-dimensional systems having z-transforms

$$\frac{b_p z^p + b_{p-1} z^{p-1} + \cdots + b_0}{z^q + a_{q-1} z^{q-1} + \cdots + a_0}$$

with bounded parameters: $|a_i| \leq K$ and $|b_i| \leq k$ for all i, and with the distance between any pole-zero pair $\geq \alpha$. $\mathfrak{M}_{fd}(p, q, K, \alpha)$ is compact in the graph topology, and on this set the graph topology is finer than the product topology.

Proof: Let $\{P_i(z)\}$ be a sequence of plants in $\mathfrak{M}_{fd}(p, q, K, \alpha)$, and suppose $P_i = N_i/D_i$, with $\deg N_i \leq p$, $\deg D_i = q$, D_i monic, and the coefficients of N_i and D_i bounded by K. Clearly, N_i and D_i lie in sets which are compact in the H_∞-topology. Hence, there exist a subsequence $N_{k_i} \to N^*$ and $D_{k_i} \to D^*$. We now verify that $P^* \equiv N^*/D^*$ is in $\mathfrak{M}_{fd}(p, q, K, \alpha)$. We first note that H_∞

[1] In this paper, the z-transform of a system with impulse response h is $\sum_{i=0}^{\infty} h_i z^i$.

convergence of polynomials of bounded degree is equivalent to convergence of their coefficients. Hence, deg $N^* \leq p$, deg $D^* = q$, D^* is monic, and their coefficients are bounded by K. Moreover, since the location of the zeros of a polynomial is continuous of its coefficients, the zeros of N_k, D_k must converge to those of N^*, D^*, respectively, and the separation between poles and zeros is maintained at a distance of at least α. Hence, $P^* \in \mathfrak{M}_{fd}(p, q, K, \alpha)$, and $P_k \to P^*$ in the graph topology. This shows that $\mathfrak{M}_{fd}(p, q, K, \alpha)$ is compact in the graph topology. Also, in $\mathfrak{M}_{fd}(p, q, K, \alpha)$, convergence in the graph implies convergence in the coefficients of the rational transfer function, which in turn implies the convergence in each component of the impulse response. This latter fact follows by inspection of the inversion formula for z-transforms. □

It is clear that the space of all finite-dimensional systems $\mathfrak{M}_{fd}$ is a countable union of sets of the form $\mathfrak{M}_{fd}(p,q,K,\alpha)$. It then follows that Theorem 3.4 can be applied on M_{fd} equipped with the gap metric, and the infinite-horizon diameter of information $D_{gap}(u, \mathfrak{M}_{fd}, \delta)$ characterizes the optimal asymptotic error $E_\infty(u, \mathfrak{M}_{fd}, \delta)$.

We shall first derive necessary conditions on the inputs u for the robustness of the asymptotic error to measurement noise, i.e., when $D_{gap}(u, \mathfrak{M}_{fd}, \delta)$ approaches 0 as δ approaches 0. This is in terms of the notion of *stability testing*: inputs $u \equiv [u^{(1)}, u^{(2)}, \cdots, u^{(N)}]$ are said to be able to *test the stability* of plants if for every unstable $h \in \mathfrak{M}_{fd}$ at least one of the inputs $u^{(i)}$ yields an unbounded output. We have the following result on the loss of robustness when the inputs are not rich enough to test stability.

Proposition 4.11: If the inputs u cannot test stability, then $D_{gap}(u, \mathfrak{M}_{fd}, \delta) = 1$ for all $\delta > 0$.

Proof: Let $\delta > 0$. Consider the infinite-horizon uncertainty set centered at the origin:

$$S_\infty(\mathfrak{M}_{fd}, u, 0, \delta) = \{g \in \mathfrak{M}_{fd}: \|u * g\|_\infty \leq \delta\}.$$

Since u cannot test the stability of plants in M_{fd}, there must be an unstable plant $h \in \mathfrak{M}_{fd}$ such that $u * h$ is bounded: by appropriate scaling, we can assume that $h \in S_\infty(\mathfrak{M}_{fd}, u, 0, \delta)$. Since the zero plant is also in this uncertainty set and the gap distance between the zero plant and any unstable plant is 1 [6], the diameter of this uncertainty set must be 1. Hence, the diameter of information, which is the diameter of the largest uncertainty set, is also 1. □

We now give explicit necessary and sufficient conditions for inputs to be able to test stability. We begin with two definitions.

Definition 4.12: For a sequence $u \in l_\infty$, let $Z(u)$ denote the set of all zeros of its z-transform $U(z)$ inside the open-unit disk. (Note that $U(z)$ is analytic inside the open-unit disk.)

Definition 4.13: A sequence u is said to excite at frequency $\omega \in [0, 2\pi]$ if

$$\limsup_{n \to \infty} \left| \sum_{k=0}^{n} u_k e^{-jk\omega} \right| = \infty$$

i.e., the Fourier series of u at ω is unbounded. Let $\Omega(u)$ denote the set of all frequences at which u excites.

We shall now give the following result, the proof of which can be found in the Appendix.

Theorem 4.14: $\mathfrak{M}_{fd}$ is testable for stability by bounded inputs $u^{(1)}, \cdots, u^{(N)}$ if and only if the inputs have the following properties:

1) $$\bigcup_{i=1}^{N} \Omega(u^{(i)}) = [0, 2\pi]$$

2) $$\bigcap_{i=1}^{N} Z(u^{(i)}) = \emptyset.$$

Hence, the inputs can test the stability of finite-dimensional plants if and only if they excite at all frequences and have no common zeros in the unit disk.

We have the following corollary.

Corollary 4.15: $\mathfrak{M}_{fd}$ is testable for stability by a single input $u \in \overline{Bl}_\infty$ if and only if u excites at all frequencies and its z-transform has no zeros inside the open-unit disk.

Neither the existence nor the nonexistence of a bounded input having both the properties required by Corollary 4.15 has been established. However, bounded inputs which excite at all frequencies do exist. In fact, Lusin [19] has constructed a sequence which excites at all frequencies despite the fact that the sequence actually tends to 0.

Stability testing is a necessary property the inputs must satisfy in order to have robustness in the asymptotic error. It will now be shown that stability testing combined with the property of containing all finite sequences of 1's and -1's are in fact sufficient to guarantee robustness.

Theorem 4.16: If the inputs u can test stability and at least one of them contains all finite sequences of 1's and -1's then for all $\delta \geq 0$,

$$D_{gap}(u, \mathfrak{M}_{fd}, \delta) \leq 2\delta.$$

Proof: Consider now the infinite-horizon uncertainty set $S_\infty(\mathfrak{M}_{fd}, u, 0, \delta)$ centered at the origin. Since all the plants in this set give zero output on the inputs and the inputs test stability, all the plants in this set must be stable. Moreover, one of the inputs contains all finite sequences of 1's and -1's. We are now in a similar situation as in Theorem 4.4, which applies to the stable plant case. Exact arguments as in the proof of that theorem show that the diameter of this uncertainty set *measured in the l_1 norm* is bounded by 2δ. Since $\mathfrak{M}_{fd}$ is balanced and convex, the diameter of information equals diameter of this set (measured in the l_1 norm). Finally, by a result proved in the Appendix, the gap distance between two plants is always bounded by the H_∞ distance, and therefore also by the l_1 distance. Hence, the diameter of information $D_{gap}(u, \mathfrak{M}_{fd}, \delta)$ measured in the gap metric is bounded by the diameter of information measured in the l_1 norm, and hence also bounded by 2δ. □

We will now exhibit two inputs which have the above desired properties. First, it will be demonstrated that any

input that contains all finite sequences of 1's and -1's excites at all frequencies.

Proposition 4.17: Let u be any sequence which contains all finite sequences of 1's and -1's. Then $\Omega(u) = [0, 2\pi]$.

Proof: Let ω_0 be an arbitrary frequency in $[0, 2\pi]$. Take any $M > 0$. The sum $\Sigma_k |\cos k\omega_0|$ is divergent, so we can find an integer L such that $\Sigma_{k=0}^{L} |\cos k\omega_0| > M$. By the definition of the sequence u, there exists an integer n_1 such that

$$(u_{n_1}, u_{n_1+1}, \cdots, u_{n_1+L}) = (1, \operatorname{sgn}(\cos\omega_0), \operatorname{sgn}(\cos 2\omega_0), \cdots, \operatorname{sgn}(\cos L\omega_0)). \tag{4.19}$$

Now,

$$\begin{aligned}\left|\sum_{k=n_1}^{n_1+L} u_k e^{-jk\omega_0}\right| &= \left|\sum_{k=n_1}^{n_1+L} \operatorname{sgn}(\cos(k-n_1)\omega_0) e^{-jk\omega_0}\right| \\ &= \left|\sum_{k=0}^{L} \operatorname{sgn}(\cos k\omega_0) e^{-jk\omega_0}\right| \\ &\geq \left|\sum_{k=0}^{L} \operatorname{sgn}(\cos k\omega_0)\cos k\omega_0\right| > M.\end{aligned}$$

This is true for every M, so $\limsup_{n\to\infty} |\Sigma_{k=0}^{n} u_k e^{-jk\omega_0}| = \infty$. □

Using two inputs, one of which contains all finite sequences of 1's and -1's and the another the unit impulse, will suffice to test stability, since the former excites at all frequencies and the latter's z-transform has no zeros in the unit disk. It follows immediately from the Theorem 4.16 that an optimal worst-case gap error of 2δ can be achieved with these two inputs.

This result shows that for finite-dimensional plants, identification in the gap metric can be performed *robust* to the noise level δ, i.e., as δ goes to zero, the identification error also goes to zero. However, we have not yet shown that the two experiments are optimal or near optimal. A lower bound to the optimal asymptotic gap error using *any* bounded inputs will now be derived. This will show that for small δ, the above experiment design is no more than a factor of two from optimality.

Proposition 4.18: For any N and inputs $u \in \overline{Bl}_\infty^N$, the optimal worst-case asymptotic gap error for finite-dimensional plants satisfies

$$E_\infty^g(u, \mathfrak{M}_{fd}, \delta) \geq \frac{\delta}{\sqrt{1+\delta^2}}.$$

Proof: To prove this result, it suffices to show that the infinite-horizon gap diameter of information satisfies

$$D_{gap}(u, \mathfrak{M}_{fd}, \delta) \geq 2\frac{\delta}{\sqrt{1+\delta^2}}.$$

We make use of the following lower bound for the gap metric [44]:

$$\delta(h, 0) \geq \frac{\|h\|_{H_\infty}}{\sqrt{1+\|h\|_{H_\infty}^2}}.$$

Now,

$$\begin{aligned} &D_{gap}(u, \mathfrak{M}_{fd}, \delta) \\ &= \sup_{h\in\mathfrak{M}_{fd}}\ \sup_{\|d\|_\infty\leq\delta} \operatorname{diam}_{gap} S_\infty(\mathfrak{M}_{fd}, u, h*u+d, \delta) \\ &\geq \operatorname{diam}_{gap} S_\infty(\mathfrak{M}_{fd}, u, 0, \delta) \\ &= \sup_{g\in\mathfrak{M}_{fd},\ \|g*u\|_\infty\leq\delta} 2\delta(g, 0) \\ &\quad \text{since } \delta(g,0) = \delta(-g,0) \\ &\geq \sup_{g\in\mathfrak{M}_{fd},\ \|g*u\|_\infty\leq\delta} 2\frac{\|g\|_{H_\infty}}{\sqrt{1+\|g\|_{H_\infty}^2}} \\ &\quad \text{using the lower bound to the gap} \\ &\geq \frac{2\delta}{\sqrt{1+\delta^2}} \\ &\quad \text{choosing } g \text{ to be an impulse with magnitude } \delta. \end{aligned}$$

□

Finally, we note that this theorem has interesting implications to identification in the closed loop. To accurately estimate the plant, it is necessary that the input satisfies the conditions in Theorem 4.14. In general it is not known whether there exists one input with that property. If not, then more information about the model set should be known. An example of such information is the knowledge of a stabilizing controller of the plant to be identified. Details on this can be found in [29], [39].

V. Conclusions

In this paper, we have approached the problem of analyzing the intrinsic limitations of identification by considering the optimal worst-case asymptotic error achievable using any input and any identification algorithm. This gives an intrinsic measure of the difficulty of identification, given the *a priori* knowledge (model set and disturbance class) and the constraints on the allowable experiments (input class).

The analysis is performed in two steps. First, for fixed inputs, a lower bound on the error of any identification algorithm is expressed in terms of the diameter of the worst-case infinite-horizon uncertainty set, and it was shown that under some compactness conditions on the model set, there exist algorithms which achieve to within a factor of two of this bound asymptotically. These results hold for any error metric and disturbance norm. Second, for specific identification problems, characterization of inputs which makes this infinite-horizon diameter of information small is given. In particular, we considered identification in both the l_1 and the H_∞ norms for stable plants, and in the gap metric for unstable finite-dimensional

plants of arbitrary order. The significance of these error metrics is that if the worst-case error is small in these metrics, methods exist for synthesizing controllers to achieve robust performance [2], [4].

The results show that accurate identification is possible in the worst case for a specific choice of inputs depending on the model set. For identification in the l_1 norm, algorithms for computing estimates are based on linear programming and are easily implementable. For the identification in the gap metric, robust identification was shown to be more or less equivalent to stability testing. This has important implications on closed-loop identification in which one does not have direct access to the input.

There are many issues in worst-case identification that need to be resolved. The issue of computational complexity and implementation of the algorithm is a central issue. In particular, it is beneficial to relate the complexity of the model set to the complexity of the required experiments and the algorithms. Another issue is the relationship between the identification in the frequency domain and the time domain, particularly as it relates to algorithms and complexity. Deeper study of identification of unstable plants in a closed-loop setting is needed. The relations of all of this to adaptive control is of course one of the prime motivations for this work and will be the subject of future research.

VI. Acknowledgments

We would like to thank Sanjeev Kulkarni, Sanjoy Mitter, and George Zames for very helpful discussions.

Appendix

A. Proof of Theorem 4.14

To prove this result, we need the following lemma, the proof of which is elementary but tedious, and can be found in [38].

Lemma A.1: Let $u \in \overline{Bl}_\infty$ and let h be a complex-valued impulse response (i.e., the sequence values can be complex) with a strictly proper rational transfer function

$$H(z) = \frac{\sum_{i=0}^{M-1} \alpha_i z^i}{(z - e^{j\omega})^M}.$$

(It has a single pole repeated M times at $e^{j\omega}$.) Then:

1) If u excites at frequency ω, the output $u * h$ is unbounded.

2) If u does not excite at ω and $M = 1$ (the pole is simple), the output $u * h$ is bounded.

Armed with this lemma, we can now prove Theorem 4.14.

Proof:

(if part)

Let $u^{(1)}, u^{(2)}, \cdots, u^{(N)} \in \overline{Bl}_\infty$ be N inputs satisfying properties (1) and (2). Let $h \in \mathfrak{M}_{fd}$ with a rational z-transform $H(z)$, and assume that the outputs $u^{(i)} * h$, $i = 1, \cdots, N$, are all bounded. We shall show that h must be stable.

Suppose that $H(z)$ has a pole $z = z_1$ inside the open-unit disk. Since the inputs have no common zeros, then one of the inputs, say $u^{(i)}$, has no zero at $z = z_1$. Hence, the output $y^{(i)}$ must have a pole at $z = z_1$, and therefore cannot be bounded.

Thus, H can only have poles on or outside of the unit circle. Write

$$H(z) = H_u(z) + H_s(z) \tag{A.20}$$

where $H_s(z)$ contains the stable poles (outside the unit circle) and the finite impulse response (FIR) part of $H(z)$, and $H_u(z)$ is strictly proper with all poles on the unit circle. Let h_u and h_s be the inverse transforms of H_u and H_s, respectively. Since the output $u * h_s$ corresponding to the stable part must be bounded, one needs only to verify that the boundedness of $u^{(i)} * h_u$ for every i implies $h_u = 0$.

Suppose that H_u is not identically 0 and has $L > 0$ poles (counting multiplicities) on the unit circle at distinct frequencies $\omega_1, \omega_2, \cdots, \omega_M$. Then $H_u(z)$ can be decomposed as

$$H_u(z) = \sum_{i=1}^{M} H_i(z) \tag{A.21}$$

where

$$H_i(z) = \frac{\sum_{k=0}^{L_i - 1} \alpha_{ik} z^k}{(z - e^{j\omega_i})^{L_i}} \tag{A.22}$$

and L_i is the order of the pole at $z = e^{j\omega_i}$.

Consider a minimal state space realization of the system with transfer function $H_u(z)$, where the states x consist of the modes corresponding to each pole of the system. The dimension of the realization is L and some of the states are complex but they occur in conjugate pairs. (These correspond to conjugate poles.) Since $\bigcup_{i=1}^{N} \Omega(u^{(i)}) = [0, 2\pi]$ the frequency ω_1 lies in $\Omega(v)$ for some input $v \in \{u^{(1)}, \cdots, u^{(N)}\}$. By Proposition A.1,

$$y^{(1)} = v * h^{(1)} \notin l_\infty \tag{A.23}$$

where $h^{(1)}$ is the impulse response whose z-transform is $H_1(z)$.

If $x^{(1)}$ are the modal states (of dimension L_1) corresponding to this pole at ω_1, the system $h^{(1)}$ can be realized minimally as

$$x_{n+1}^{(1)} = A_1 x_n^{(1)} + B_1 v_n, \qquad y_n^{(1)} = C_1 x_n^{(1)} \tag{A.24}$$

for some matrices A_1, B_1, C_1.

Since $y^{(1)}$ is unbounded but v is bounded, it follows from (A.24) that the modal states $x^{(1)}$ must be unbounded given input v. But the overall state x for the entire system $H_u(z)$ is an aggregation of the modal states and hence must become unbounded too when input v is applied. The last step is to show that this implies that the output of the overall system must be unbounded also.

Let the minimal state space realization of H_u be

$$x_{n+1} = Ax_n + Bv_n, \qquad y_n = Cx_n. \tag{A.25}$$

From (A.25), a sequence of equations is obtained as

$$y_n = Cx_n$$
$$y_{n+1} = CAx_n + CBv_n$$
$$\vdots$$
$$y_{n+L-1} = CA^{L-1}x_n + \sum_{i=0}^{L-2} CA^iBv_n.$$

Let

$$\mathbf{y}_n = \begin{bmatrix} y_n \\ y_{n+1} \\ \vdots \\ y_{n+L-1} \end{bmatrix}, \qquad Q_0(A,C) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{L-1} \end{bmatrix},$$

$$E = \begin{bmatrix} 0 \\ CB \\ \vdots \\ \sum_{i=0}^{L-2} CA^iB \end{bmatrix}.$$

The sequence of output equations can then be written as

$$\mathbf{y}_n = Q_0(A,C)x_n + Ev_n. \tag{A.26}$$

Note that $Q_0(A,C)$ is the observability matrix of the system by the minimality of the realization, $Q_0(A,C)$ is invertible. Since x_n becomes unbounded and v_n is bounded, the output $\mathbf{y}_n$ must be also unbounded. This contradicts our original assumption and hence $H_u \equiv 0$. The original system h must be stable and the inputs $u^{(1)},\cdots,u^{(N)}$ can test stability in $\mathfrak{M}_{fd}$.

(only-if part)

We now show that the two conditions for the inputs are also necessary to test the stability in $\mathfrak{M}_{fd}$.

Suppose the first condition is not satisfied; consider an $\omega_0 \in [0,2\pi]$ but $\omega_0 \notin \cup_{i=1}^{N} \Omega(u^{(i)})$. Consider the unstable system $h_n = \cos(n\omega_0)$. Lemma A.1(b) implies that $u^{(i)} * e^{jn\omega_0}$ is bounded for all i. Since $u^{(i)} * h$ is the real part of $u^{(i)} * e^{jn\omega_0}$, it is also bounded for all i. Thus, the inputs cannot test stability in $\mathfrak{M}_{fd}$. This shows that the first condition is necessary.

Now suppose that the second condition is not satisfied, so that there exists some $z_0 = r_0e^{j\omega_0}$ $(0 < r_0 < 1)$ which is a common zero in the open-unit disk of the z-transforms of all the inputs: that is,

$$\sum_{k=0}^{\infty} u_k^{(i)} r_0^k e^{jk\omega_0} = 0, \qquad \forall i. \tag{A.27}$$

Since the inputs are real, their zeros occur as conjugate pairs, i.e.,

$$\sum_{k=0}^{\infty} u_k^{(i)} r_0^k e^{-jk\omega_0} = 0 \qquad \forall i. \tag{A.28}$$

Now consider the unstable finite-dimensional system $h_n = r_0^{-n}\cos(n\omega_0)$. For each i,n,

$$\begin{aligned}
|(u^{(i)} * h)_n| &= \left|\sum_{k=0}^{n} u_k^{(i)} r_0^{-(n-k)} \cos(n-k)\omega_0\right| \\
&= \left|\frac{1}{2} r_0^{-n} \sum_{k=0}^{n} u_k^{(i)} r_0^k \left(e^{j(n-k)\omega_0} + e^{-j(n-k)\omega_0}\right)\right| \\
&= \frac{1}{2} r_0^{-n} \left| e^{jn\omega_0}\left(\sum_{k=0}^{n} u_k^{(i)} r_0^k e^{-jk\omega_0}\right)\right. \\
&\qquad \left. + e^{-jn\omega_0}\left(\sum_{k=0}^{n} u_k^{(i)} r_0^k e^{jk\omega}\right)\right| \\
&= \frac{1}{2} r_0^{-n} \left| e^{jn\omega_0}\left(-\sum_{k=n+1}^{\infty} u_k^{(i)} r_0^k e^{-jk\omega_0}\right)\right. \\
&\qquad \left. + e^{-jn\omega_0}\left(-\sum_{k=n+1}^{\infty} u_k^{(i)} r_0^k e^{jk\omega_0}\right)\right| \\
&\le r_0^{-n} \sum_{k=n+1}^{\infty} r_0^k = \frac{r}{1-r}.
\end{aligned}$$

Thus the output for each of the inputs is bounded. Hence, the inputs $u^{(i)}$'s cannot test the stability in $\mathfrak{M}_{fd}$. □

B. An Inequality Between the Gap and H_∞ Distances

Proposition B.1: Let h and g be two plants. Then

$$\delta(g,h) \le \|h-g\|_{H_\infty}.$$

Proof: We assume that $\|h-g\|_{H_\infty} < \infty$ otherwise there is nothing to prove. Now,

$$\delta(g,h) = \max\left(\vec{\delta}(G_g,G_h), \vec{\delta}(G_h,G_g)\right)$$

where

$$G_h \equiv \{(u, h*u) \in l_2 : x \in l_2, h*x \in l_2\}$$

and

$$\vec{\delta}(G_h,G_g) \equiv \sup_{x\in G_h, \|x\|_2\le 1} \inf_{y\in G_g} \|x-y\|_2.$$

Now, since $\|g-h\|_{H_\infty} < \infty$

$$(u, h*u) \in G_h \Leftrightarrow (u, g*u) \in G_g.$$

We have

$$\begin{aligned}
\vec{\delta}(G_h,G_g) &\le \sup_{h*u\in l_2, \|u\|_2\le 1} \inf_{y\in G_g} \|(u,h*u)-y\|_2 \\
&\le \sup_{h*u\in l_2, \|u\|_2\le 1} \|(u,h*u)-(u,g*u)\|_2 \\
&= \sup_{h*u\in l_2, \|u\|_2\le 1} \|(h-g)*u\|_2 \\
&\le \|h-g\|_{H_\infty}.
\end{aligned}$$

Hence, the result follows. □

References

[1] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. Warmuth, "Occam's razor," *Inf. Proc. Lett.*, vol. 24, pp. 377–380, 1987.

[2] M. A. Dahleh and M. H. Khammash. "Controller design in the presence of structured uncertainty." to appear in *Automatica* special issue on robust control.

[3] M. A. Dahleh. T. Theodosopoulos. and J. N. Tsitsiklis. "The sample complexity of worst-case identification of F. I. R. linear systems." *Syst. Contr. Lett.*. to be published.

[4] J. C. Doyle. "Analysis of feedback systems with structured uncertainty." *IEE Proc.*. vol. 129. pp. 242–250. 1982.

[5] E. Fogel and Y. F. Huang. "On the value of information in system identification-bounded noise case." *Automatica*. vol. 18. no. 2. pp. 229–238. 1982.

[6] T. T. Georgiou and M. C. Smith. "Optimal robustness in the gap metric." *IEEE Trans. Automat. Contr.*, vol. 35. no. 6. pp. 673–686. 1990.

[7] G. C. Goodwin. "Experiment design in system identification" in *Encyclopedia of Systems and Control*. M. Singh. Ed. New York: Pergamon Press. 1987.

[8] G. Gu and P. P. Khargonekar. "Linear and nonlinear algorithms for identification in $\mathcal{H}_\infty$ with error bounds. *IEEE Trans. Automat. Contr.*, vol. 37. no. 7. July 1992.

[9] A. J. Helmicki. C. A. Jacobson. and C. N. Nett. "Identification in H_∞: A robust convergent nonlinear algorithm." in *Proc. 1989 Int. Symp. Math. Theory Networks Syst.*, 1989.

[10] ——, "Identification in H_∞: Linear algorithms." in *Proc. 1990 Amer. Contr. Contr.*, pp. 2418–2423.

[11] A. J. Helmicki. C. A. Jacobson. and C. N. Nett. "Control-oriented system identification: A worst-case deterministic approach in H_∞," *IEEE Trans. Automat. Contr.*. vol. 36. no. 10. Oct. 1991.

[12] C. A. Jacobson and C. N. Nett. "Worst-case system identification in l_1: Optimal algorithms and error bounds." in *Proc. 1991 Amer. Contr. Conf.*, June 1991.

[13] B. Kacewicz and M. Milanese. "On the optimal experiment design in the worst-case l_1 *system identification.*" submitted to CDC.

[14] B. Kacewicz and L. Plaskota. "Noisy information for linear systems in asymptotic setting," *J. Complexity*, vol. 7. pp. 35–57. 1991.

[15] J. M. Krause. G. Stein. and P. P. Khargonekar. "Robust performance of adaptive controllers with general uncertainty structure." in *Proc. 29th Conf. Decision Contr.*. 1990. pp. 3168–3175.

[16] R. Lozano-Leal and R. Ortega. "Reformulation of the parameter identification problem for systems with bounded disturbances." *Automatica*. vol. 23. no. 2. pp. 247–251. 1987.

[17] M. K. Lau. R. L. Kosut. and S. Boyd. "Parameter set estimation of systems with uncertain nonparametric dynamics and disturbances." in *Proc. 29th Conf. Decision Contr.*. 1990. pp. 3162–3167.

[18] L. Lin. L. Wang, and G. Zames. "Uncertainty principle and identification n-widths for LTI and slowly varying systems." ACC. Chicago. IL. 1992.

[19] A. Markushevich. *Theory of Functions of A Complex Variable*. Englewood Cliffs. NJ: Prentice-Hall. vol. I. 1965.

[20] R. K. Mehra. "Optimal input signals for parameter estimation in dynamic systems—A survey and new results." *IEEE Trans. Automat. Contr.*. vol. AC-19. pp. 753–768. 1974.

[21] C. A. Michelli and T. J. Rivlin. "A survey of optimal recovery" in *Optimal Estimation in Approximation Theory*. (C. A. Michelli and T. J. Rivlin. Eds. New York: Plenum. 1977.

[22] M. Milanese and G. Belforte. "Estimation theory and uncertainty intervals evaluation in the presence of unknown but bounded errors: Linear families of models and estimators." *IEEE Trans. Automat. Contr.*. vol. AC-27. pp. 408–414. 1982.

[23] M. Milanese and R. Tempo. "Optimal algorithm theory for robust estimation and prediction." *IEEE Trans. Automat. Contr.*, vol. AC-30. pp. 730–738. 1985.

[24] M. Milanese. "Estimation theory and prediction in the presence of unknown and bounded uncertainty: A survey." in *Robustness in Identification and Control*. M. Milanese. R. Tempo. A. Vicino. Eds. New York: Plenum Press. 1989.

[25] P. M. Makila. "Robust identification and galois sequences." *Int. J. Contr.*. vol. 54. pp. 1189–1200.

[26] P. M. Makila and J. R. Partington. "Robust approximation and identification in H_∞," in *Proc. 1991 Amer. Cont. Conf.*, June, 1991.

[27] J. Munkres. *Topology—A First Course*. Englewood Cliffs. NJ: Prentice-Hall. 1975.

[28] J. P. Norton. "Identification and application of bounded-parameter models." *Automatica*. vol. 23. no. 4, pp. 497–507, 1987.

[29] J. R. Partington and P. M. Makila. "Worst-case identification of stabilizable systems." submitted.

[30] J. Pearl. "On the connection between the complexity and credibility of inferred models." *Int. J. General Syst.*, vol. 4, pp. 255–264. 1978.

[31] K. Poolla and A. Tikku. "On the time complexity of worst-case system identification." preprint. 1992.

[32] L. Pronzato and E. Walter. "Experiment design in bounded-error context: Comparison with D-optimality," *Automatica*. vol. 25. no. 3. pp. 383–391. 1989.

[33] A. K. El-Sakkary. "The Gap Metric: Robustness of Stabilization of Feedback Systems." *IEEE Trans. Automat. Contr.*, vol. AC-30. pp. 240–247. 1985.

[34] T. Theodosopoulos. "Worst-case identification in l_1: Algorithms and complexity," thesis in preparation.

[35] J. F. Traub and H. Wozniakowski. *A General Theory of Optimal Algorithms*. New York: Academic. 1980.

[36] J. F. Traub. G. Wasilkowski. and H. Wazniakowski. *Information-Based Complexity*. New York: Academic. 1988.

[37] D. N. C. Tse. M. A. Dahleh. and J. N. Tsitsiklis. "Optimal and Robust Identification in the l_1 norm." in *Proc. 1991 Amer. Contr. Conf.*, June 1991.

[38] D. N. C. Tse. "Optimal and robust identification under bounded disturbances." master's thesis. Dept. Elec. Eng. Comp. Sci.. M.I.T.. Feb. 1991.

[39] D. N. C. Tse. M. A. Dahleh. and J. N. Tsitsiklis. "Optimal asymptotic worst-case identification with applications on l_1 and gap metrics." in *Proc. Int. Symp. MTNS-91*. pp. 325–330.

[40] M. Vidyasagar. *Control System Synthesis*. Cambridge. MA: M.I.T. Press. 1985.

[41] G. Zames. "On the metric complexity of casual linear systems: ϵ-entropy and ϵ-dimension for continuous-time." *IEEE Trans. Automat. Contr.*. vol. 24. Apr. 1979.

[42] G. Zames and A. El-Sakkary, "Unstable systems and the gap metric." in *Proc. Allerton Conf.*, pp. 380–385, Oct. 1980.

[43] M. Zarrop. *Optimal Experimental Design for Dynamic System Identification*. New York: Springer-Verlag, 1979.

[44] S. Zhu. M. Hautus. and C. Praagman. "A lower and upper bound for the gap metric." in *Proc. 28th Conf. Decision Contr.*, Dec. 1989. pp. 2337–2341.

David N. C. Tse (S'91) was born in Hong Kong. in 1966. He received the B.A.Sc. degree in systems design engineering from University of Waterloo. Canada. in 1989. and the M.S. degree in electrical engineering from Massachusetts Institute of Technology, Cambridge, MA. in 1991.

He is currently a Ph.D. student in the Laboratory for Information and Decision Systems at MIT. His research interests include data compression. stochastic processes. system identification. and information theory.

Munther A. Dahleh (S'84–M'87) was born in August 1962. He received the B.S.E.E. degree from Texas A & M University, College Station. TX in 1983. and the Ph.D. degree from Rice University, Houston. TX 1987.

Since 1987. he has been with the Department of Electrical Engineering and Computer Science. Massachusetts Institute of Technology, Cambridge, MA. where he is currently an Associate Professor. In July 1992 he was awarded the Finmeccanica career development chair. He has held consulting positions with NASA and the C.S. Draper Laboratory since 1988. and with FIAT Research Center since 1991. His current interests include robust control, identification of uncertain systems and adaptation. and control of time-varying and infinite-dimensional systems.

Dr. Dahleh is currently serving as an Associate Editor for the *Systems and Control Letters*.

John N. Tsitsiklis (S'80–S'81–M'81–M'83) was born in Thessaloniki, Greece, in 1958. He received the B.S. degree in mathematics, in 1980, and the B.S., M.S., and Ph.D. degrees in electrical engineering, in 1980, 1981, and 1984, respectively, all from the Massachusetts Institute of Technology, Cambridge, MA.

During the academic year 1983–1984, he was an acting Assistant Professor of Electrical Engineering at Stanford University, Stanford, CA. Since 1984, he has been with the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, where he is currently an Associate Professor. His research interests are in the areas of systems and control theory, parallel and distributed computation, and operations research. He has coauthored more than 50 papers in these areas. He is also a coauthor (with D. Bertsekas) of *Parallel and Distributed Computation: Numerical Methods*, 1989. He has been a recipient of an IBM Faculty Development Award, 1983, an NSF Presidential Young Investigator Award, 1986, an Outstanding Paper Award by the IEEE Control Systems Society (for a paper coauthored with M. Athans, 1986), and of the Edgerton Faculty Achievement Award by M.I.T., 1989. He has been a plenary speaker at the *1992 IEEE Conference on Decision and Control*. Finally, he is an Associate Editor of *Automatica* and *Applied Mathematics Letters*, and has been an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.

Minimization of the Maximum Peak-to-Peak Gain: The General Multiblock Problem

Ignacio J. Diaz-Bobillo and Munther A. Dahleh

***Abstract*—This paper presents a comprehensive study of the general l_1-optimal multiblock problem, as well as a new linear programming algorithm for computing suboptimal controllers. By formulating the interpolation conditions in a concise and natural way, the general theory is developed in simpler terms and with a minimum number of assumptions. In addition, further insight is gained on the structure of the optimal solution, and different classes of multiblock problems are distinguished. This leads to conceptually attractive, iterative method for finding approximate solutions with the following properties: 1) approximates multiblock problems with one-block problems by delay augmentation, 2) unifies the treatment of zero and rank interpolation conditions through robust computations, 3) provides upper and lower bounds of the optimal objective function by solving one finite dimensional linear program at each iteration, 4) for a class of problems, it generates suboptimal controllers that achieve the upper bound without order inflation, 5) both bounds as well as the solution converge to the optimal, 6) it does not require the existence of polynomial feasible solutions, and 7) gives information about the support structure of the optimal solution.**

Notation

Let X be a real normed vector space, then X^* denotes the dual space of X containing all bounded linear functionals on X.

- l_1 Space of absolutely summable sequences supported on the nonnegative integers. If $x \in l_1$ then $\|x\|_1 = \sum_{k=0}^{\infty}|x(k)| < \infty$.
- $l_1^{p\times q}$ Space of $p \times q$ matrices with entries in l_1. If $M = (m_{ij}) \in l_1^{p\times q}$, then $\|M\|_1 := \max_{1\le i\le p}\sum_{j=1}^{q}\|m_{ij}\|_1$.
- l_∞ Space of all bounded sequences of real numbers supported on the nonnegative integers. If $x \in l^\infty$ then $\|x\|_\infty := \sup_k|x(k)| < \infty$.
- $l_\infty^{p\times q}$ Space of $p \times q$ matrices with entries in l_∞. If $M = (m_{ij}) \in l_\infty^{p\times q}$, then $\|M\|_\infty := \sum_{i=1}^{p}\max_{1\le j\le q}\|m_{ij}\|_\infty$. Note that $l_\infty^{p\times q} = (l_1^{p\times q})^*$.
- $c_0^{p\times q}$ Subspace of $l_\infty^{p\times q}$ consisting of all elements whose entries decay to zero, i.e., $\lim_{k\to\infty}m_{ij}(k) = 0$ for all $\{ij\}$. Note that $(c_0^{p\times q})^* = l_1^{q\times q}$.
- λ Complex variable representing the unit delay. Given $M \in l_1^{p\times q}$, define $\hat{M}(\lambda) := \sum_{k=0}^{\infty}M(k)\lambda^k$ as the λ-transform of M.
- $\mathscr{D}$ The open unit disk.
- P_k The truncation operator on sequences. Hence, if $x = \{x(i)\}_{i=0}^{\infty}$ is any sequence, then $P_k x = \{x(0), x(1), \cdots, x(k), 0, \cdots\}$.
- S_k Right shift by k positions. If $x = \{x(i)\}_{i=0}^{\infty}$ is any sequence and k is a nonnegative integer, then

$$S_k x = \{\overbrace{0, \cdots, 0}^{k}, x(0), x(1), \cdots\}.$$

Given a matrix M, $(M)_i$ will denote its ith row and $(M)^j$ its jth column.

Manuscript received March 23, 1992; revised November 20, 1992. Paper recommended by Past Associate Editor, T. T. Georgiou. This work was supported by NSF under grant 9157306-ECS, by C. S. Draper Laboratory under grant DL-H-441636, and by AFOSR under grant 0368.

The authors are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 9211898.

I. Introduction

DESIGN specifications for practical control problems are often most naturally expressed in terms of time-domain bounds on the amplitude of signals (exogenous disturbances and regulated outputs). This observation has led to the introduction of a new optimization problem in the context of control system design. In [37] Vidyasagar formulated the l_1-optimal control problem. In contrast with the $\mathscr{H}_\infty$ problem, the l_1-optimal design has as objective the minimization of the maximum peak-to-peak gain of a closed-loop system that is driven by bounded amplitude disturbances.

From 1987–1988, Dahleh and Pearson introduced some basic results on the theory of l_1 optimization. In [9] the solution to the l_1-optimal control problem was presented for the special case of square (i.e., one-block) systems. Then, in [11] Dahleh *et al.* presented the central ideas for the solution of nonsquare (i.e., multiblock) problems, including a method to compute approximate suboptimal solutions iteratively. Such method is based on the solution of a linear program representing a truncated version of the original problem. Similar results extending these ideas to the continuous-time domain were introduced by the same authors in [10], as well as a solution to the fixed input optimization problem [12].

These results brought considerable attention to the problem of l_1 optimization. In [29] a general treatment of the multiblock case was presented, where the optimal solution is shown to exits under some assumptions. Independently in [6] and [33] a method was introduced to compute lower bounds on the optimal norm, by solving a complementary linear program. A direct linear programming formulation (in the primal space) was presented in [30]. Also, [34] introduced a nice account of some convergence properties and pointed to interesting deficiencies in

0018-9286/93$03.00

the theory. In [17], [18] the full state-feedback problem was addressed.

On the area of robustness, considerable advancement was made too. In [13], the necessity of the small gain theorem in the l_1 context was analyzed. Also, [24] presented necessary and sufficient conditions for robust performance and robust stability under structured time-varying perturbations. It turns out that such conditions are relatively easy to compute making the theory more attractive from the point of view of applications. Other related work can be found in [8], [6], [3], [19], [14], [32].

The present investigation is motivated by the lack of a solid understanding of the general l_1 multiblock problem. While various aspects of the theory are well understood, the structure of the optimal solution in the general multiblock case is not. As a result, solution methods which are based on a straightforward truncation of the full problem, suffer from significant deficiencies. Most important, they generate a sequence of suboptimal controllers of increasing order, and miss the structure of the (possibly low order) optimal controller. This issue was pointed out quite nicely in [33] where exact solutions of low order were computed for some example. From a practical point of view, such truncation method translates into high order controllers even for the simplest multiblock problems. At the same time, it requires the existence of feasible closed-loop maps with finite pulse response, a condition that many control problems lack.

In this paper we present a comprehensive treatment of the general l_1-optimal multiblock problem. Contributions are made in the general theory as well as in the approximate methods of solution. With regard to the problem formulation, a more compact and natural way of characterizing the interpolation conditions of the general multiblock problem is presented. It has the advantage of simplifying many of the proofs and avoiding unnecessary assumptions (compared to previous work [29], [34]). We also present a new solution method for the general multiblock problem with the following characteristics:

1) Approximates multiblock problems with one-block problems by delay augmentation, thus exploiting the characteristics of the optimal solutions of such problems.

2) Applies results from matrix theory [21] in the computation of interpolation conditions.

3) With each approximation (requiring the solution of only one linear program), the method provides upper and lower bounds of the optimal norm.

4) Under mild assumptions, both bounds converge to the optimal value of the norm.

5) With each approximation the method generates a feasible (i.e., stabilizing) controller that achieves the upper bound.

6) For a special class of multiblock problems the solutions are exact.

7) For a larger class of multiblock problems the sequence of suboptimal controllers does not suffer from order inflation.

Also, a result is presented relating the support characteristics of the optimal and approximate solution of multiblock problems, followed by a stronger conjecture. These results are complemented by a broad range of numerical examples, including a case study where the l_1 and $\mathscr{H}_\infty$ solution to the pitch axis control of the X29 aircraft are compared.

The paper is organized as follows: in Section II the general l_1-optimal control problem is defined. The new interpolation conditions are presented in Section III as well as computational procedures. This is followed by an existence result with minimum assumptions in Section IV. Next, we establish the equivalence between l_1 optimization and infinite dimensional linear programming in Section V. Section VI contains the solution to one-block problems. The results in this section are an extension of those in [29]. Section VII presents (approximate) methods of solution to multiblock problems. In particular, the delay augmentation method is introduced along with its convergence properties. Illustrations and examples are contained in Section VIII. In Sections IX and X, we present a few results and observations (including a conjecture) on the support characteristics of these approximate solutions. Finally, we treat the X29 synthesis problem in Section XI followed by the conclusions in Section XII.

II. Problem Formulation

The setup corresponds to the standard disturbance rejection problem formulated as a linear fractional transformation from the disturbance input to the regulated output, with the controller in the lower loop (see Fig. 1). In particular, we consider the discrete time case, with the inputs and outputs being sequences of vectors. The problem is represented via an LTI finite-dimensional operator, G, that maps the disturbance vector w of dimension n_w, and the control vector u of dimension n_u, to the regulated output vector z of dimension n_z, and the measurement vector y of dimension n_y. Thus, with the appropriate partitioning,

$$\begin{pmatrix} z \\ y \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \begin{pmatrix} w \\ u \end{pmatrix}. \tag{1}$$

The controller action is represented by the operator K that maps the measurement sequence to the control sequence, i.e., $u = Ky$. The closed-loop map from the disturbance to the regulated output, denoted Φ, is given by:

$$\Phi = G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21}. \tag{2}$$

The l_1-optimal control problem can be stated as follows: among all internally stabilizing controllers, find the one that minimizes the maximum peak-to-peak gain of Φ operating on the space of bounded disturbances with unit norm. That is,

$$\begin{aligned} \mu^o &:= \inf_{K \text{ stab}} \sup_{\max_{1 \le i \le n_w} \|w_i\|_\infty = 1} \left(\max_{1 \le k \le n_z} \|(\Phi w)_k\|_\infty \right) \\ &= \inf_{K \text{ stab}} \|\Phi\|_1. \end{aligned} \tag{3}$$

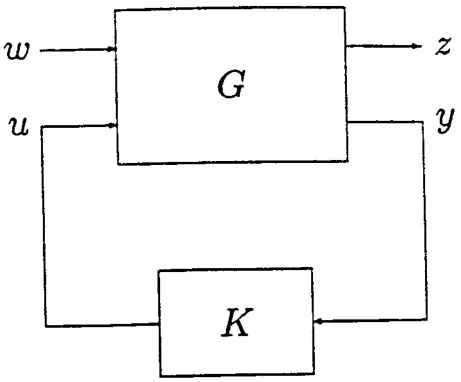


Fig. 1. The standard problem.

In the above we have used the fact that the induced norm of an operator mapping bounded sequences in $\mathbb{R}^{n_w}$ to bounded sequences in $\mathbb{R}^{n_z}$ is given by the $l_1^{n_z\times n_w}$ norm.

It is well known that a simpler description of the set of all (internally) stable closed-loop maps is obtained via a parameterization of all stabilizing controllers [38]. Such parameterization provides an affine expression, mapping an operator space to the set of all internally stable closed-loop maps:

$$\Phi = H - UQV \tag{4}$$

where $H \in l_1^{n_z\times n_w}$, $U \in l_1^{n_z\times n_u}$ and $V \in l_1^{n_y\times n_w}$ are functions of the problem data (i.e., the operator G), and Q is a free parameter in $l_1^{n_u\times n_y}$ (i.e., stable). Furthermore, if G is LTI and finite dimensional, so are H, U, and V. Then, for any $Q \in l_1^{n_u\times n_y}$, a controller can be computed that achieves the corresponding closed-loop map, Φ.

Consequently, the l_1 problem can be redefined as a minimum distance problem in $l_1^{n_z\times n_w}$:

$$\mu^o := \inf_{R\in\mathscr{S}} \|H - R\|_1 = \inf_{\Phi - H\in\mathscr{S}} \|\Phi\|_1 \tag{5}$$

where

$$\mathscr{S} := \{R \in l_1^{n_z\times n_w} | R = UQV \text{ for some } Q \in l_1^{n_u\times n_y}\}. \tag{6}$$

The subspace $\mathscr{S}$ contains the set of feasible R's. Also, from duality theory [26], problem (5) can be posed in the dual space of $l_1^{n_z\times n_w}$, that is, $l_\infty^{n_z\times n_w}$ as the following maximization problem:

$$\mu^o = \max_{\substack{G\in\mathscr{S}^\perp\\ \|G\|_\infty\le 1}} \langle H, G\rangle \tag{7}$$

where $\langle H, G\rangle$ is the value of the bounded linear functional G at the point H:

$$\langle H, G\rangle = \sum_{i=1}^{n_z}\sum_{j=1}^{n_w}\sum_{k=0}^{\infty} g_{ij}(k)h_{ij}(k)$$

and $\mathscr{S}^\perp$ is the right annihilator of $\mathscr{S}$:

$$\mathscr{S}^\perp = \{G \in l_\infty^{n_z\times n_w} | \langle R, G\rangle = 0 \quad \forall R \in \mathscr{S}\}.$$

Furthermore, if a solution to (5) exists, say Φ^o, then it is aligned with every solution G^o to (7), that is $\langle \Phi^o, G^o\rangle = \|\Phi^o\|_1\|G^o\|_\infty$. This implies that Φ^o and G^o must satisfy the following alignment conditions:

i) if $|g_{ij}^o(t)| < \max_{1\le j\le n_w}\|g_{ij}^o\|_\infty$, then $\phi_{ij}^o(t) = 0$,

ii) $\phi_{ij}^o(t)g_{ij}^o(t) \ge 0$,

iii) let $I = \{i \in [1,2,\cdots,n_z] | (G^o)_i \equiv 0\}$, then $\|(\Phi^o)_i\|_1 = \mu^o$ for all i not in I,

iv) for all $i \in I$, $(\Phi^o)_i$ can be anything such that $\|(\Phi^o)_i\|_1 \le \mu^o$.

The next section studies the solvability of the equation $R = UQV$ for Q in $l_1^{n_u\times n_y}$.

III. Interpolation Conditions

Here we take some of the ideas in [11] and [29], and present a natural and compact description of the interpolation conditions for the most general MIMO case.

The notion of interpolation conditions can be viewed in at least two ways: as algebraic conditions on the matrix $\hat{R}(\lambda)$ so that it belongs to the range of $\hat{U}\hat{Q}\hat{V}$, or as conditions on the nullspace of the operator R. Here we are going to exploit the algebraic notion although, for the purpose of computations, we view the interpolation conditions as a nullspace matching problem.

In the sequel it will be assumed, without loss of generality, that $\hat{U}(\lambda)$ has full column rank (i.e., rank of n_u for almost all λ) and $\hat{V}(\lambda)$ has full row rank (i.e., rank of n_y for almost all λ). Violation of these assumptions implies that there are redundancies in the controls and/or the measurements which can be easily removed.

First, a simple but useful result from complex variable theory is presented, where $(\cdot)^{(k)}(\lambda_0)$ denotes the kth order derivative with respect to λ, evaluated at λ_0.

Lemma 3.1: Given a function $f(\cdot)$ of the complex variable λ analytic in $\mathscr{D}$, then $(f)^{(k)}(\lambda_0) = 0$ for $k = 0, 1, \cdots, (\sigma - 1)$ for $\lambda_0 \in \mathscr{D}$ if and only if $f(\lambda) = (\lambda - \lambda_0)^\sigma g(\lambda)$ where $g(\cdot)$ is analytic in $\mathscr{D}$.

Next, consider Smith–McMillan decompositions of the rational matrices $\hat{U}$ and $\hat{V}$. (Note: to simplify notation, the complex variable argument will be omitted in most expressions)

$$\hat{U} = \hat{L}_U\hat{M}_U\hat{R}_U \tag{8}$$

$$\hat{V} = \hat{L}_V\hat{M}_V\hat{R}_V \tag{9}$$

where $\hat{L}_U$, $\hat{R}_U$, $\hat{L}_V$, and $\hat{R}_V$ are (polynomial) unimodular matrices. Under the rank assumptions on $\hat{U}$ and $\hat{V}$, the rational matrices $\hat{M}_U$ and $\hat{M}_V$ have the following diagonal structure:

$$\hat{M}_U = \begin{pmatrix} \frac{\epsilon_1}{\psi_1} & & \\ & \ddots & \\ & & \frac{\epsilon_{n_u}}{\psi_{n_u}} \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \tag{10}$$

$$\hat{M}_V = \begin{pmatrix} \frac{\epsilon_1'}{\psi_1'} & & & 0 & \cdots & 0 \\ & \ddots & & \vdots & \ddots & \vdots \\ & & \frac{\epsilon_{n_y}'}{\psi_{n_y}'} & 0 & \cdots & 0 \end{pmatrix}. \tag{11}$$

Let λ_0 be a zero of $\hat{U}(\lambda)$. Let $\sigma_{U_i}(\lambda_0)$ denote the multiplicity of λ_0 as a root of $\epsilon_i(\lambda)$, then $\{\sigma_{U_i}(\lambda_0)\}_{i=1}^{n_u}$ defines a nondecreasing sequence of nonnegative integers. For a given $i \in \{1, 2, \cdots, n_u\}$, $\sigma_{U_i}(\lambda_0)$ is known as the algebraic multiplicity of λ_0. The total number of indexes i for which $\sigma_{U_i}(\lambda_0)$ is strictly positive is known as the geometric multiplicity of λ_0. Similarly, define $\{\sigma_{V_j}(\lambda_0)\}_{j=1}^{n_y}$ for $\hat{V}(\lambda)$.

Let Λ_{UV} denote the set of zeros of $\hat{U}$ and $\hat{V}$ in $\bar{\mathscr{D}}$. In order to prove the interpolation theorem (i.e., apply the results of Lemma 3.1) we need the following assumption.

Assumption 1: $\Lambda_{UV} \subset \mathscr{D}$.

Consider the unimodular matrices in (8). Since their inverses are polynomials, one can define the following polynomial row and column vectors:

$$\hat{\alpha}_i(\lambda) = \left(\hat{L}_U^{-1}\right)_i(\lambda) \qquad i = 1, 2, \cdots, n_z$$

$$\hat{\beta}_j(\lambda) = \left(\hat{R}_V^{-1}\right)^j(\lambda) \qquad j = 1, 2, \cdots, n_w. \tag{12}$$

Now we are ready to present the main interpolation theorem. These conditions are different from those in [29] and do not require coprime factorizations.

Theorem 3.1: Given $R \in l_1^{n_z \times n_w}$, there exists $Q \in l_1^{n_u \times n_y}$ such that $R = UQV$ if and only if for all $\lambda_0 \in \Lambda_{UV} \subset \mathscr{D}$ the following conditions are satisfied:

i) $(\hat{\alpha}_i \hat{R} \hat{\beta}_j)^{(k)}(\lambda_0) = 0$ for

$$\cdot \begin{cases} i = 1, \cdots, n_u \\ j = 1, \cdots, n_y \\ k = 0, \cdots, \sigma_{U_i}(\lambda_0) + \sigma_{V_j}(\lambda_0) - 1 \end{cases}$$

$$\text{ii)} \begin{cases} \left(\hat{\alpha}_i \hat{R}\right)(\lambda) \equiv 0 & \text{for } i = n_u + 1, \cdots, n_z \\ \left(\hat{R} \hat{\beta}_j\right)(\lambda) \equiv 0 & \text{for } j = n_y + 1, \cdots, n_w. \end{cases}$$

Proof: Consider the following factorization of $\hat{M}_U$ and $\hat{M}_V$ (where $\mathbf{0}$ denotes a block of zeros of appropriate dimensions):

$$\hat{M}_U =: \begin{pmatrix} \hat{\mathscr{E}}_U \hat{\Psi}_U^{-1} \\ \mathbf{0} \end{pmatrix}; \qquad \hat{M}_V =: \begin{pmatrix} \hat{\Psi}_V^{-1} \hat{\mathscr{E}}_V & \mathbf{0} \end{pmatrix}$$

where $\hat{\mathscr{E}}_U$ and $\hat{\mathscr{E}}_V$ retain the zeros in Λ_{UV} while $\hat{\Psi}_U$ and $\hat{\Psi}_V$ capture the stable (i.e., minimum-phase) zeros of $\hat{U}$ and $\hat{V}$ along with their (stable) poles. Thus, both Ψ_U and Ψ_V are invertible in l_1. Then,

$$\hat{R} = \hat{L}_U \begin{pmatrix} \hat{\mathscr{E}}_U \hat{\bar{Q}} \hat{\mathscr{E}}_V & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \hat{R}_V$$

where $\hat{\bar{Q}} := \hat{\Psi}_U^{-1} \hat{R}_U \hat{Q} \hat{L}_V \hat{\Psi}_V^{-1}$. Clearly, $\bar{Q} \in l_1^{n_u \times n_y}$ if and only if $Q \in l_1^{n_u \times n_y}$. Next, define the following partitions of $\hat{L}_U$ and $\hat{R}_V$:

$$\hat{L}_U = \begin{pmatrix} \hat{L}_{U,1} & \hat{L}_{U,2} \end{pmatrix}; \qquad \hat{R}_V = \begin{pmatrix} \hat{R}_{V,1} \\ \hat{R}_{V,2} \end{pmatrix} \tag{13}$$

where $\hat{L}_{U,1}$ has n_u columns and $\hat{R}_{V,1}$ has n_y rows. Then, given $R \in l_1^{n_z \times n_w}$,

$$\exists Q \in l_1^{n_u \times n_y} \text{ such that } R = UQV$$
$$\Updownarrow$$
$$\exists \bar{Q} \in l_1^{n_u \times n_y} \text{ such that } R = L_{U,1} \mathscr{E}_U \bar{Q} \mathscr{E}_V R_{V,1}.$$

Necessity of condition *i*) follows immediately. Take any $i \in \{1, \cdots, n_u\}$ and $j \in \{1, \cdots, n_y\}$, then

$$\left(\hat{\alpha}_i \hat{R} \hat{\beta}_j\right)(\lambda) = \prod_{\lambda_0 \in \Lambda_{UV}} (\lambda - \lambda_0)^{\sigma_{U_i}(\lambda_0)} \hat{\bar{q}}_{ij}(\lambda) \cdot \prod_{\lambda_0 \in \Lambda_{UV}} (\lambda - \lambda_0)^{\sigma_{V_j}(\lambda_0)}$$

which implies condition *i*) by Lemma 3.1 and the fact that $\bar{q}_{ij}$ is in l_1.

Necessity of condition *ii*) results from the following: take any $i \in \{n_u + 1, \cdots, n_z\}$ and $j \in \{n_y + 1, \cdots, n_w\}$, then $(\hat{\alpha}_i \hat{R})(\lambda) \equiv 0$ and $(\hat{R} \hat{\beta}_j)(\lambda) \equiv 0$ since $(\hat{\alpha}_i \hat{L}_{U,1})(\lambda) \equiv 0$ and $(\hat{R}_{V,1} \hat{\beta}_j)(\lambda) \equiv 0$.

To show that conditions *i*) and *ii*) are sufficient we proceed by backwards construction: by Lemma 3.1,

$$\text{i)} \Rightarrow \begin{pmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_{n_u} \end{pmatrix} \hat{R} \begin{pmatrix} \hat{\beta}_1 & \cdots & \hat{\beta}_{n_y} \end{pmatrix} = \hat{\mathscr{E}}_U \hat{W} \hat{\mathscr{E}}_V$$

for some $W \in l_1^{n_u \times n_y}$ since $R \in l_1^{n_z \times n_w}$. Moreover,

$$\text{ii)} \Rightarrow \begin{pmatrix} \hat{\alpha}_{n_u+1} \\ \vdots \\ \hat{\alpha}_{n_z} \end{pmatrix} \hat{R} \equiv \mathbf{0} \quad \text{and} \quad \hat{R} \begin{pmatrix} \hat{\beta}_{n_y+1} & \cdots & \hat{\beta}_{n_w} \end{pmatrix} \equiv \mathbf{0}.$$

Therefore, combining these equations into one,

$$\hat{L}_U^{-1} \hat{R} \hat{R}_V^{-1} = \begin{pmatrix} \hat{\mathscr{E}}_U \hat{W} \hat{\mathscr{E}}_V & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

which implies that $W = \bar{Q}$ is the solution. ■

In words, Theorem 3.1 provides a set of algebraic conditions which are necessary and sufficient for R to be feasible (i.e., equivalent to UQV for some stable Q). The conditions in i) make sure that the left and right unstable zero structure of the composition UQV is preserved while the conditions in ii) impose the correct (normal) rank conditions on $\hat{R}$. In fact, it is possible to view the collection of $\hat{\alpha}_i$'s and $\hat{\beta}_j$'s for $i > n_u$ and $j > n_y$, as two polynomial basis (not necessarily of minimal degree) for the left and right nullspaces of $\hat{R}(\lambda)$ (see [23]). By virtue of the Smith–McMillan decomposition these sets of polynomial vectors are linearly independent (over the field of rational functions) so they generate a minimal set of constraints on $\hat{R}$ (Note: the four-block case has some

redundancy which can be eliminated *a priori*, see [18] for a detailed discussion).

In the sequel, we refer to the conditions in i) as the *zero interpolation conditions*, and to the conditions in ii) as the *rank interpolation conditions*. Rank interpolation conditions are also known by the names of relations [11] and convolution conditions [33], [34].

Problems of the form (4) have been traditionally classified in the $\mathscr{H}_\infty$ and $\mathscr{H}_2$ literature according to the dimensions of the different signal spaces involved. Here we adopt the same classification:

- *One-Block Problems*: When $n_w = n_y$ and $n_z = n_u$. These are also known as good rank or square problems.
- *Two-Block Column Problems*: When $n_w = n_y$ and $n_z > n_u$.
- *Two-Block Row Problems*: When $n_w > n_y$ and $n_z = n_u$.
- *Four-Block Problems*: When $n_w > n_y$ and $n_z > n_u$.

A problem is labeled multiblock when it is not one-block. Multiblock problems are also known as bad rank problems [11], [29].

Clearly, one-block problems only require zero interpolation conditions and have no rank interpolation conditions, while two-block row (column) problems require right (left) rank interpolation conditions, and four-block problems require both left and right rank interpolation conditions.

A. Computation of Interpolation Conditions

The problem of finding the Smith–McMillan decomposition of rational matrices is at the heart of the interpolation problem. This decomposition has been studied thoroughly due to its strong connections with several important notions in system theory (e.g., multivariable zeros and poles), although mostly from an algebraic point of view [23]. The standard algebraic algorithm to compute such objects is based on the Euclidean division algorithm, known to be numerically sensitive. Nevertheless, there has been some effort in this direction, for example, by using symbolic methods from computer algebra on polynomial matrices [4]. However, it is generally desirable to have algorithms based on the state-space representation of systems, that are more easily implemented on digital computers.

Here we present an alternative approach to the problem of finding the zero interpolation conditions of a square rational matrix. Such approach avoids the explicit computation of the Smith–McMillan decomposition. Furthermore, it is computationally attractive since it is based on finding the nullspaces of certain Toeplitz-like matrices which are formed directly from the state-space representation of the system.

Although multiblock problems require rank interpolation conditions, we will show that those problem can be posed in such a way that only zero interpolations need to be considered.

In Theorem 3.1 we have shown how the internal stability of the closed-loop system is assured if the zero structure of the left unstable zeros of $\hat{U}$ and the right unstable zero of $\hat{V}$ is preserved in $\hat{R}$. Such structure is characterized by the zero frequency, its algebraic and geometric multiplicity, and its directional properties as given by the corresponding polynomial vector $\hat{\alpha}_i$ or $\hat{\beta}_j$. Despite its numerical problems, the Smith–McMillan decomposition provides the most natural way of characterizing the zero and pole structure of a rational matrix. To circumvent the formal Smith–McMillan decomposition of $\hat{U}(\lambda)$ and $\hat{V}(\lambda)$, it is necessary to find an alternative set of conditions that unequivocally defines the zero structure of a rational matrix. Such a set is presented in this section.

The theory of zeros of MIMO systems has been studied extensively, both from an algebraic and state-space perspective [28], [16], [31]. It is well known that a zero of a square system given in state-space form $[A, B, C, D]$, is characterized by the solution of a generalized eigenvalue problem of the form [28]:

$$\begin{pmatrix} A - z_0 I & B \\ C & D \end{pmatrix} \begin{pmatrix} x_0 \\ u_0 \end{pmatrix} = 0$$

where $z_0 := \lambda_0^{-1}$, x_0 is known as the state zero direction and u_0 is known as the zero input direction. However, the numerical stability of such eigenvalue problem deteriorates quickly when there are zeros with algebraic multiplicity greater than one. Indeed, such difficulty is equivalent to finding the Jordan decomposition of a defective matrix (i.e., a nondiagonalizable matrix) which is known to be a hard numerical problem [22].

Although it is diffcult to obtain the full zero structure directly from the state-space description of a system, the location or frequency of the zeros can be reliably computed [20]. In the sequel, we will assume that the locations of the unstable zeros of the rational (square) matrices $\hat{U}(\lambda)$ and $\hat{V}(\lambda)$ are available.

Following, we introduce a useful definition along with some notation.

Definition 3.1: Given a rational matrix $\hat{H}(\lambda)$ analytic at λ_0 and a positive integer σ; define the following block-lower-triangular Toeplitz matrix:

$$T_{\lambda_0,\sigma}(\hat{H}) = \begin{pmatrix} H_0 & 0 & 0 & \cdots & 0 \\ H_1 & H_0 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ H_{\sigma-1} & H_{\sigma-2} & H_{\sigma-3} & \cdots & H_o \end{pmatrix} \tag{14}$$

where the H_i's are given by the Taylor expansion of $\hat{H}(\lambda)$ at λ_0, that is,

$$\hat{H}(\lambda) = H_0 + (\lambda - \lambda_0)H_1 + (\lambda - \lambda_0)^2 H_2 + (\lambda - \lambda_0)^3 H_3 + \cdots$$

and $H_i = (1/i!)(\hat{H})^{(i)}(\lambda_0)$.

A numerically stable method was proposed in [36] to find the structural indices associated with poles and zeros

of a stable rational matrix $\hat{H}$, by looking at the rank of $T_{\lambda_0,\sigma}(\hat{H})$ as σ increases. Such approach, however, does not provide the directional information necessary to construct the interpolation conditions. Here we present an extension of the ideas in [36] by looking at the structure of the nullspace of $T_{\lambda_0,\sigma}(\hat{H})$ for increasing values of σ. Such approach has strong connections with the general interpolation theory of rational matrix functions [1], [2]. In particular, it exploits the analyticity of the matrices $\hat{U}$ and $\hat{V}$ in the disk.

The following definition establishes some terminology [1].

Definition 3.2: Given an $m \times n$ (real) rational matrix $\hat{H}(\lambda)$ analytic at λ_0, a right null chain of order σ at λ_0 is an ordered set of column vectors in $\mathbb{R}^n$, $\{x_1, x_2, \cdots, x_\sigma\}$, such that $x_1 \neq 0$ and

$$T_{\lambda_0,\sigma}(\hat{H})\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_\sigma \end{pmatrix} = \mathbf{0}.$$

Similarly, a left null chain of order σ at λ_0 is an ordered set of column vectors in $\mathbb{R}^m$, $\{y_1, y_2, \cdots, y_\sigma\}$, such that $y_1 \neq 0$ and

$$T_{\lambda_0,\sigma}(\hat{H}^T)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_\sigma \end{pmatrix} = \mathbf{0}.$$

The next Theorem shows that, if $\hat{H}$ is square, the existence of a right (left) null chain of order σ at λ_0 is equivalent to the existence of a zero at λ_0 of algebraic multiplicity σ. It is an extension of [21, theorem 1.12]. Later, we will establish a complete equivalence between the structure of a zero and the null chains associated with that zero.

Theorem 3.2: A full rank, $n \times n$, rational matrix $\hat{H}(\lambda)$, analytic at λ_0, has a zero at λ_0 of geometric multiplicity l and a sequence of structural indexes equal to, at least, $\sigma_{n-l+1}, \cdots, \sigma_n$ $(\sigma_1 = \cdots = \sigma_{n-l} = 0)$ if and only if the following conditions hold

1) There exist l polynomial vectors, $\hat{u}_1, \cdots, \hat{u}_l$, such that

$$\left(\hat{H}\hat{u}_j\right)^{(k)}(\lambda_0) = 0 \quad \text{for} \quad k = 0, \cdots, \sigma_{n-l+j} - 1 \qquad \forall j = 1, \cdots, l.$$

2) The set of vectors $\{\hat{u}_1(\lambda_0), \cdots, \hat{u}_l(\lambda_0)\}$ is linearly independent and

$$\operatorname{span}\{\hat{u}_1(\lambda_0), \cdots, \hat{u}_l(\lambda_0)\} = \mathcal{N}\left[\hat{H}(\lambda_0)\right].$$

where $\mathcal{N}[\cdot]$ denotes the null space of a matrix.

Proof: Necessity follows directly from the Smith–McMillan decomposition of $\hat{H}(\lambda)$:

$$\hat{H}(\lambda) = \hat{L}(\lambda)\hat{M}(\lambda)\hat{R}(\lambda).$$

Say that the jth entry $(j \geq n - l + 1)$ on the diagonal of $\hat{M}$ has a factor $(\lambda - \lambda_0)^{\sigma_j}$. Then, pick $\hat{u}_{j-n+l}$ to be the jth column of $\hat{R}^{-1}$. With this choice

$$\begin{aligned} \hat{H}\hat{u}_{j-n+l} &= \hat{H}(\hat{R}^{-1})^j \\ &= (\lambda - \lambda_0)^{\sigma_j}\hat{L}\hat{p}_{j-n+l} \qquad \forall j = n - l + 1, \cdots, n \end{aligned}$$

where $\hat{p}_{j-n+l}(\lambda)$ is a rational vector analytic at λ_0. Clearly, this implies that $(\hat{H}\hat{u}_{j-n+l})^{(k)}(\lambda_0) = 0$ for $k = 0, \cdots, \sigma_j - 1$, and further the set $\{\hat{u}_1(\lambda_0), \cdots, \hat{u}_l(\lambda_0)\}$ is linearly independent since $\hat{R}$ is unimodular and spans the null space of $\hat{H}(\lambda_0)$.

The proof of sufficiency is not as straightforward. Let $\hat{z}_j := \hat{H}\hat{u}_j$ $j = 1, \cdots, l$ and define the following auxiliary rational vectors:

$$\hat{y}_j(\lambda) := \left(\hat{L}^{-1}\hat{z}_j\right)(\lambda), \quad \hat{v}_j(\lambda) := \left(\hat{R}\hat{u}_j\right)(\lambda) \qquad j = 1, \cdots, l.$$

Then, we have that $\hat{y}_j(\lambda) = \hat{M}(\lambda)\hat{v}_j(\lambda)$. Note that $\hat{u}_1(\lambda_0) \cdots \hat{u}_l(\lambda_0)$ are linearly independent if and only if $\hat{v}_1(\lambda_0) \cdots \hat{v}_l(\lambda_0)$ are linearly independent since R is unimodular. Further, since multiplication by a unimodular matrix preserves the zero structure, this direction of the proof can be restated as follows: let $j = 1, \ldots, l$, then

$$\begin{gathered} \exists \hat{v}_j(\lambda) \text{ such that } \hat{v}_1(\lambda_0) \cdots \hat{v}_l(\lambda_0) \text{ are linearly} \\ \text{independent and } \hat{y}_j^{(k)}(\lambda_0) = 0,\ k = 0, \cdots, \sigma_{n-l+j-1} \\ \Downarrow \\ \exists (\lambda - \lambda_0)^{\sigma_{n-l+j}} \text{ in the } n - l + j \text{ diagonal entry of } \hat{M}(\lambda). \end{gathered}$$

Now, it follows from above that

$$\hat{y}_j(\lambda) = (\lambda - \lambda_0)^{\sigma_{n-l+j}}\hat{p}_j(\lambda).$$

Let $\hat{\epsilon}_j(\lambda)$, $j = 1, \cdots, n$ be the diagonal entries of the matrix $\hat{M}$. It immediately follows that:

$$\begin{aligned} &\begin{pmatrix} \hat{\epsilon}_1 & & \\ & \ddots & \\ & & \hat{\epsilon}_n \end{pmatrix}(\hat{v}_1(\lambda) \cdots \hat{v}_l(\lambda)) \\ &= (\hat{p}_1(\lambda) \cdots \hat{p}_l(\lambda))\begin{pmatrix} (\lambda - \lambda_0)^{\sigma_{n-l+1}} & & \\ & \ddots & \\ & & (\lambda - \lambda_0)^{\sigma_n} \end{pmatrix}. \end{aligned} \tag{15}$$

First, we show that the matrix $(\hat{v}_1(\lambda_0) \cdots \hat{v}_l(\lambda_0))$ has the structure

$$\begin{pmatrix} 0 \\ \hat{V}(\lambda_0) \end{pmatrix}. \tag{16}$$

The top zero block results from the fact that the matrix $\hat{M}(\lambda_0)$ has a null space of dimension l (otherwise there will be more linearly independent vectors than l), hence, $\hat{\epsilon}_1, \cdots, \hat{\epsilon}_{n-l}$ do not have zeros at λ_0. From (15), it follows

that for all λ

$$\begin{pmatrix} \hat{\epsilon}_{n-l+1} & & \\ & \ddots & \\ & & \hat{\epsilon}_n \end{pmatrix} \hat{V} = \hat{P} \begin{pmatrix} (\lambda-\lambda_0)^{\sigma_{n-l+1}} & & \\ & \ddots & \\ & & (\lambda-\lambda_0)^{\sigma_n} \end{pmatrix}. \tag{17}$$

where the matrices $\hat{V}$ and $\hat{P}$ are obtained from the decompositions

$$(\hat{v}_1(\lambda) \cdots \hat{v}_l(\lambda)) = \begin{pmatrix} \hat{\bar{V}} \\ \hat{V} \end{pmatrix}$$

and

$$(\hat{p}_1(\lambda) \cdots \hat{p}_l(\lambda)) = \begin{pmatrix} \hat{\bar{P}} \\ \hat{P} \end{pmatrix}.$$

Let

$$\hat{E} = \begin{pmatrix} \hat{\epsilon}_{n-l+1} & & \\ & \ddots & \\ & & \hat{\epsilon}_n \end{pmatrix},$$

$$\hat{D} = \begin{pmatrix} (\lambda-\lambda_0)^{\sigma_{n-l+1}} & & \\ & \ddots & \\ & & (\lambda-\lambda_0)^{\sigma_n} \end{pmatrix}.$$

Then, from (16), it is clear that $\hat{V}(\lambda_0)$ has full rank. Let $\hat{R}_1$, $\hat{R}_2$ be unimodular matrices such that

$$\hat{V}\hat{R}_1 = \hat{L} \quad \text{where } \hat{L} \text{ is lower triangular}$$

and

$$\hat{R}_2\hat{P} = \hat{U} \quad \text{where } \hat{U} \text{ is upper triangular}.$$

From this (17) can be factored as follows:

$$\hat{E}\hat{L} = \hat{R}_2^{-1}\hat{U}\hat{D}\hat{R}_1.$$

Clearly, the matrix $\hat{E}\hat{L}$ has the same zero structure as the matrix $\hat{U}\hat{D}$. By direct computation of the Smith matrix of $\hat{U}\hat{D}$, it follows that $(\lambda-\lambda_0)^{\sigma_{n-l+j}}$ is a factor of jth diagonal element. Since $\hat{L}$ has full rank at λ_0, it follows that $(\lambda-\lambda_0)^{\sigma_{n-l+j}}$ is a factor of $\hat{\epsilon}_{n-l+j}$. This completes the proof. ∎

Note that a similar result holds for left zeros simply by replacing $\hat{H}$ and $\hat{H}^T$. The following corollary restates the result of Theorem 3.2 in terms of null chains.

Corollary 3.1: A full rank, square, rational matrix $\hat{H}(\lambda)$ analytic at λ_0, has a right (left) zero at λ_0 of (at least) algebraic multiplicity σ if and only if there exits a right (left) null chain of order σ at λ_0.

Proof: Both directions of the proof follow immediately by equating

$$\hat{u}(\lambda) = x_1 + (\lambda-\lambda_0)x_2 + \cdots + (\lambda-\lambda_0)^{\sigma-1}x_\sigma.$$

∎

Note that if $\hat{H}$ has a right zero of geometric multiplicity greater than one, say l, then there are l different right null chains (not necessarily of the same order), such that the span of the x_1's equals the nullspace of $\hat{H}(\lambda_0)$. Let $x^i(y^i)$ denote the ith right (left) null chain of order σ_i, then the following definition applies [1].

Definition 3.3: A canonical set of right null chains of $\hat{H}(\lambda)$ at λ_0 is an ordered set of right null chains, i.e., $x^i = (x_1^i \cdots x_{\sigma_i}^i)$ for $i = 1,\cdots,l$, such that

i) $\{x_1^1, x_1^2,\cdots,x_1^l\}$ are linearly independent,

ii) $\operatorname{span}\{x_1^1, x_1^2,\cdots,x_1^l\} = \mathcal{N}[\hat{H}(\lambda_0)]$, and

iii) $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_l$.

A canonical set of left null chains is defined similarly.

Next, we show that the zero interpolation conditions of Theorem 3.1 can be stated in terms of the canonical set of right null chains of $\hat{V}$ and the canonical set of left null chains of $\hat{U}$ at each $\lambda_0 \in \Lambda_{UV}$. For that we need to introduce an extension of the above definition.

Definition 3.4: An extended set of right null chains of a full rank $n \times n$ rational matrix $\hat{H}(\lambda)$ at λ_0, is a canonical set of right null chains augmented with $n-l$ vectors in $\mathbb{R}^n$, i.e., $\{x_1^{l+1},\cdots,x_1^n\}$, such that the span of $\{x_1^1, x_1^2,\cdots,x_1^n\}$ is equal to $\mathbb{R}^n$. The order associated with these added chains is zero.

From the above definition, if a square rational matrix has no zeros at λ_0, then the corresponding canonical set of null chains is empty and the extended set is a basis for $\mathbb{R}^n$, e.g., the columns of an $n \times n$ identity matrix.

Next, we apply the above results and definitions to the zero interpolation conditions of a one-block problem. In the context of Theorem 3.1 we have the following equivalence: for $j = 1,\cdots,n_y$ and $k = 0,\cdots,\sigma_{V_j}-1$,

$$\left(\hat{V}\hat{\beta}_j\right)^{(k)}(\lambda_0) = 0 \Leftrightarrow T_{\lambda_0,\sigma_{V_j}}(\hat{V})x^{n_y-j+1} = 0$$

where x^i is an extended set of right null chains for $\hat{V}$ at λ_0. The sequence of x^i's has to be reversed in the above equation due to the fact that σ_{V_j} is a nondecreasing sequence of algebraic multiplicities while an extended set of null chains is defined with the opposite ordering. Note that if $\sigma_{V_j} = 0$ then both conditions are satisfied trivially (i.e., there are no conditions). Similarly, for $i = 1,\cdots,n_u$ and $k = 0,\cdots,\sigma_{U_i}-1$.

$$\left(\hat{\alpha}_i\hat{U}\right)^{(k)}(\lambda_0) = 0 \Leftrightarrow T_{\lambda_0,\sigma_{U_i}}(\hat{U}^T)y^{n_u-i+1} = 0.$$

In other words, the extended set of left and right null chains are locally (i.e., for each λ_0) equivalent to the polynomial vectors $\hat{\alpha}_i$'s and $\hat{\beta}_j$'s. Having made this observation, we are ready to present an alternative set of zero interpolation conditions.

Given an element of an extended set of right null chains at λ_0, x^j, of order σ_j, define the following polyno-

mial vector:

$$\hat{x}^j_{\lambda_0}(\lambda) := x^j_1 + (\lambda - \lambda_0)x^j_2 + \cdots + (\lambda - \lambda_0)^{\sigma_j - 1}x^j_{\sigma_j}$$

if $\sigma_j > 0$, and $\hat{x}^j_{\lambda_0}(\lambda) := x^j_1$ if $\sigma_j = 0$. Similarly, define $\hat{y}^i_{\lambda_0}(\lambda)$ for an element of an extended set of left null chains, y^i, of order σ_i. With this notation we have the following corollary.

Corollary 3.2: Given a one-block problem, the zero interpolation conditions of Theorem 3.1 are equivalent to the following: for all $\lambda_0 \in \Lambda_{UV}$,

$$\left(\hat{y}^i_{\lambda_0}\hat{R}\hat{x}^j_{\lambda_0}\right)^{(k)}(\lambda_0) = 0$$

$$\text{for} \begin{cases} i = 1,\cdots,n_u \\ j = 1,\cdots,n_y \\ k = 0,\cdots,\sigma_{U_i}(\lambda_0) + \sigma_{V_j}(\lambda_0) - 1 \end{cases}$$

where y^i and x^j are elements of the extended sets of left and right null chains of $\hat{U}$ and $\hat{V}$ respectively, and σ_{U_i} and σ_{V_j} are the corresponding orders (i.e., algebraic multiplicities).

Proof: Follows directly from Theorems 3.1 and 3.2, and from the above definitions. ■

B. Computation of Null Chains

This subsection discusses a simple algorithm to compute the extended set of null chains at λ_0 of a full rank square rational matrix analytic at λ_0. Let $\hat{H}(\lambda)$ denote an $n \times n$ rational matrix and assume that λ_0 is given, then the algorithm is based on the computation of a basis for the nullspace of $T_{\lambda_0,\sigma}(\hat{H})$ for increasing values of σ.

Consider the construction of an extended set of right null chains. By Definition 3.2, given some positive integer σ, any vector in the kernel of $T_{\lambda_0,\sigma}(\hat{H})$ such that $x_1 \neq 0$ is a potential member of the set. Let B_σ denote a matrix whose columns form a basis for the right nullspace of $T_{\lambda_0,\sigma}(\hat{H})$, then the following algorithm generates an extended set of right null chains.

Step 1) Compute B_σ for $\sigma = 1, 2, \cdots$ until the top n rows are filled with zeros (no more null chains can be extracted at this point). Then the maximum order of any chain, σ_1, is given by the current value of the counter (σ) minus one. Note that, by Corollary 3.1, this iteration process is guaranteed to stop since the rational matrix $\hat{H}$ is finite dimensional (i.e., its zeros have finite algebraic multiplicity).

Step 2) Let b_i for $i = 1,\cdots,r$ denote each column of B_{σ_1}. Reduce the dimension of the b_i's by removing all sets of n contiguous zeros at the top of each vector. The result is a collection of r vectors (possibly of different dimensions) such that the top n entries of each one define a nonzero vector in $\mathbb{R}^n$. (Note that at least one will have dimension $n\sigma_1$.)

Step 3) Sort the resulting vectors in decreasing order of dimension. Let l be the rank of the $n \times r$ matrix that results from collecting the first n rows of each vector. Then, select the first l vectors such that the reduced matrix that results from collecting the first n rows of each vector has rank l. Such collection forms a canonical set of right null chains.

Step 4) Extend the set by augmenting the collection with $n - l$ vectors such that the set of n vectors formed with the first n rows define a basis in $\mathbb{R}^n$.

If the system $\hat{H}(\lambda)$ is given in state-space form, say $[A, B, C, D]$, then the Toeplitz matrices $T_{\lambda_0,\sigma}(\hat{H})$ can be easily computed using the following equation (see Definition 3.1):

$$H_k = \begin{cases} \lambda_0 C(I - \lambda_0 A)^{-1}B + D & \text{for } k = 0 \\ C(I - \lambda_0 A)^{-k-1}A^{k-1}B & \text{for } k = 1, 2, \cdots. \end{cases}$$

Note that $(I - \lambda_0 A)^{-1}$ always exists since λ_0 is in the unit disk and $\hat{H}$ is stable (i.e., analytic in the closed unit disk). A word of warning is necessary, however, when λ_0 is close to the unit circle and A has a stable eigenvalue that is also close to the unit circle and next to λ_0. Such cases may give rise to numerical difficulties. Besides this fact, the rest of the algorithm only involves the computation of nullspaces that can be done efficiently through the well known QR or singular value decompositions [22].

C. A Simple Example

In order to illustrate the workings of the algorithm introduced in the previous section, a simple example is presented. Let $\hat{H}(\lambda)$ be a 3×3 polynomial matrix given by:

$$\hat{H}(\lambda) = \begin{pmatrix} (\lambda - 0.5)^2 & \lambda(\lambda + 2)(\lambda - 0.5) & 0 \\ (\lambda - 0.5)^3 & \lambda(\lambda - 0.5) & 0 \\ 0 & 0 & \lambda^2 \end{pmatrix}.$$

We have chosen a polynomial matrix just to make the example tractable without the aid of a computer. Let us construct an extended set of right null chains for the zero at $\lambda_0 = 0.5$. According to step one, we compute the nullspace of $T_{\lambda_0,\sigma}(\hat{H})$ for $\sigma = 1, 2, \cdots$. In particular, for $\sigma = 3$ we have:

$$T_{0.5,3}(\hat{H}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & .25 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & .25 & 0 & 0 & 0 \\ 1 & 1.5 & 0 & 0 & .5 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & .5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & .25 \end{pmatrix};$$

$$B_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Clearly, the first three rows of B_3 are zero so we stop increasing σ. Then, the maximum algebraic multiplicity of $\lambda_0 = 0.5$ is two, i.e., $\sigma_1 = 2$. Next, (Step 2), reduce each column of B_3 by eliminating the leading blocks of zeros to get:

$$b_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}; \quad b_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}; \quad b_3 = \begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}.$$

Then (Step 3), reorder the set of vectors in decreasing dimension, i.e., $\{b_3, b_1, b_2\}$, and compute the rank of the matrix formed with the first three rows:

$$l = \operatorname{rank}\left[\begin{pmatrix}1 & 1 & 0\\0 & 0 & 1\\0 & 0 & 0\end{pmatrix}\right] = 2.$$

Then, the canonical set of right null chains is given by $\{x^1, x^2\}$ where

$$x^1 = \begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix} \quad \text{and} \quad x^2 = \begin{pmatrix}0\\1\\0\end{pmatrix}$$

with their corresponding order (i.e., algebraic multiplicity) being $\sigma_1 = 2$ and $\sigma_2 = 1$. This indicates that the geometric multiplicity of λ_0 is two. Finally (Step 4), to get an extended set of right null chains we augment the collection with $x^3 = (0\ 0\ 1)^T$ having order $\sigma_3 = 0$ (by definition).

IV. Duality and Existence

With Theorem 3.1 we have established a compact algebraic characterization of the set $\mathcal{S}$. Next, we need to interpret these results in the context of (7), which calls for the identification of the subspace of $l_\infty^{n_z \times n_w}$ which annihilates $\mathcal{S}$.

Following the approach in [11] and [29], we write the zero interpolation conditions as functionals acting on R. Indeed, for all (i, j, k) in the ranges established in Theorem 3.1, for $l = 0, 1, \cdots$, and all $\lambda_0 \in \Lambda_{UV}$, define $RF_{ijk\lambda_0}$ and $IF_{ijk\lambda_0}$ in $l_\infty^{n_z \times n_w}$ such that

$$\left[RF_{ijk\lambda_0}(l)\right]_{qp} := \sum_{t=0}^{\infty}\sum_{s=0}^{\infty} \alpha_{iq}(s-l) \cdot \beta_{pj}(t-s)\,\Re\left[(\lambda^t)^{(k)}\right]\Big|_{\lambda=\lambda_0} \tag{18}$$

and

$$\left[IF_{ijk\lambda_0}(l)\right]_{qp} := \sum_{t=0}^{\infty}\sum_{s=0}^{\infty} \alpha_{iq}(s-l) \cdot \beta_{pj}(t-s)\,\Im\left[(\lambda^t)^{(k)}\right]\Big|_{\lambda=\lambda_0} \tag{19}$$

where $\Re(\lambda)$ and $\Im(\lambda)$ denote the real and imaginary part of λ respectively, and α_{iq} denotes the qth column of α_i while β_{pj} denotes the pth row of β_j. By straightforward algebra it can be shown that $\langle R, RF_{ijk\lambda_0}\rangle = 0$ and $\langle R, IF_{ijk\lambda_0}\rangle = 0$ if and only if $\hat{R}$ satisfies the zero interpolation conditions of Theorem 3.1. Note that only a finite number of sequences are required, thus the subspace spanned by the sequences associated with the zero interpolations is finite dimensional. In fact, the number of functionals is given by:

$$c_z := \sum_{\lambda_0 \in \Lambda_{UV}} \sum_{i=1}^{n_y} \sum_{j=1}^{n_u} \sigma_{U_i}(\lambda_0) + \sigma_{V_j}(\lambda_0). \tag{20}$$

A note should be made on the way c_z is computed. If a given $\lambda_0 \in \Lambda_{UV}$ is complex then $\bar{\lambda}_0 \in \Lambda_{UV}$ too, since $\hat{U}$ and $\hat{V}$ are real-rational. However, for the purpose of constructing functionals, only one of each pair of complex-conjugate zeros should be considered since the other one would generate redundant functionals. But, for the purpose of counting the number of independent functionals (i.e., computing c_z), both zeros should be included in Λ_{UV}, since a complex-conjugate pair of zeros generate twice as many functionals as a real zero.

Next, we look at the rank interpolation conditions [i.e., conditions in *ii*)]. Again, these algebraic conditions can be viewed as convolution of sequences. For $i = n_u + 1, \cdots, n_z$ and $q = 1, \cdots, n_w$, define the following sequence of $n_z \times n_w$ matrices:

$$X_{\alpha_i q t}(l) := \begin{pmatrix} \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \alpha_i^T(t-l) & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \end{pmatrix} \quad (\alpha_i^T(t-l) \text{ in the } q\text{th column}) \tag{21}$$

where $t, l \in \mathbf{Z}_+$. Similarly, for $j = n_y + 1, \cdots, n_w$ and $p = 1, \cdots, n_z$, define

$$X_{\beta_j p t}(l) := \begin{pmatrix} \cdots & 0 & \cdots \\ & \vdots & \\ \cdots & 0 & \cdots \\ \cdots & \beta_j^T(t-l) & \cdots \\ \cdots & 0 & \cdots \\ & \vdots & \\ \cdots & 0 & \cdots \end{pmatrix} \begin{matrix} \\ \\ \\ \}p\text{th row}. \\ \\ \\ \\ \end{matrix} \tag{22}$$

Then, $\langle R, X_{\alpha_i q t}\rangle = 0$ and $\langle R, X_{\beta_j p t}\rangle = 0$ for $t = 0, 1, \cdots$ if and only if $\hat{R}$ satisfies the rank interpolation conditions of Theorem 3.1. Note that, in contrast with the zero interpolation sequences, the linear span of the $X_{\alpha_i q t}$'s and $X_{\beta_j p t}$'s is infinite dimensional since for every (i, q, p), t can take infinite values (i.e., $t \in \mathbf{Z}_+$).

The next theorem gives a sufficient condition for the existence of an optimal solution to (5). The proof is omitted since the arguments involved are essentially the same as those in [11], [29].

Theorem 4.1: If every $\lambda_0 \in \Lambda_{UV}$ is strictly inside the unit disk, then there exists $R^o \in \mathscr{S}$ such that

$$\mu^o = \|H - R^o\|_1 = \inf_{R \in \mathscr{S}} \|H - R\|_1.$$

Note, however, that the above result is more general than that in [29], where it is assumed that $\hat{U}$ and $\hat{V}$ have square partitions with no zeros on the unit circle. Such extra assumption was avoided by determining the full set of interpolation conditions directly from the Smith–McMillan decomposition of $\hat{U}$ and $\hat{V}$.

V. l_1 Optimization and Linear Programming

This section will establish the equivalence between the primal-dual pair of optimization problems (5)–(7) and a primal-dual pair of infinite dimensional linear programs.

By definition, $\mathscr{S}^{\perp} \subset l_{\infty}^{n_z \times n_w}$ is the linear span of the sequences (18)–(22), and G is any element in that subspace with infinity norm not greater than one. That is,

$$G \in \operatorname{span}\left\{RF_{ijk\lambda_0}, IF_{ijk\lambda_0}, X_{\alpha,qt}, X_{\beta,pt}\right\} \tag{23}$$

with the appropriate index ranges.

In order to bring (5) and (7) into a standard linear programming form, it is convenient to redefine the notation, the purpose being to express both the objective and the feasible subspace in (infinite) matrix form. This is possible since the constraints that specify the feasible subspace $\mathscr{S}$ are no more and no less than an infinite collection of linear functionals annihilating the sequence R, which can be expressed as an infinite collection of equality constraints on the elements of the sequence Φ.

To bring the primal objective function $\|\Phi\|_1$ into linear form and avoid the nonlinearity built into the one norm, we use a standard change of variables from linear programming: let $\Phi = \Phi^+ - \Phi^-$, where Φ^+ and Φ^- are sequences of $n_z \times n_w$ matrices with nonnegative entries. That is, with a slight abuse of notation, $\Phi^+ \geq 0$ and $\Phi^- \geq 0$. Then, the l_1 norm of Φ takes the form $\max_i \sum_{j=1}^{n_w} \sum_{t=0}^{\infty} (\phi_{ij}^+(t) + \phi_{ij}^-(t))$ which is linear in (Φ^+, Φ^-). This expression holds only if, for any (i, j, t), either $\phi_{ij}^+(t)$ or $\phi_{ij}^-(t)$ is zero, which is a guaranteed property of the optimal solution. Indeed, if a feasible solution is such that $\phi_{ij}^+(t)$ and $\phi_{ij}^-(t)$ are strictly positive, then reducing both variables by $\min(\phi_{ij}^+(t), \phi_{ij}^-(t))$ reduces the value of the cost and does not violate feasibility since the difference remains the same, and further, one of the two variables becomes zero. Therefore, the optimal solution will always be such that either $(\phi_{ij}^+(t)$ or $\phi_{ij}^-(t))$ is zero. Note that this transformation doubles the number of variables representing the closed-loop response.

Consequently, the primal problem (5) can be restated as follows:

$$\mu^o = \inf_{\mu, \Phi^+, \Phi^-} \mu$$

subject to

$$\sum_{j=1}^{n_w} \sum_{t=0}^{\infty} \left(\phi_{ij}^+(t) + \phi_{ij}^-(t)\right) \leq \mu \qquad \text{for } i = 1, \cdots, n_z \tag{24}$$

$$\Phi - H \in \mathscr{S}.$$

Next, we shift attention to the linear constraints representing the feasible set. From the previous discussion it is clear that a given Φ is feasible (i.e., there exists a stable Q such that $\Phi = H - UQV$) if and only if

$$\begin{cases} \langle \Phi, RF_{ijk\lambda_0} \rangle = \langle H, RF_{ijk\lambda_0} \rangle \\ \langle \Phi, IF_{ijk\lambda_0} \rangle = \langle H, IF_{ijk\lambda_0} \rangle \end{cases}$$

$$\text{for} \begin{cases} \lambda_0 \in \Lambda_{UV} \\ i = 1, \cdots, n_u \\ j = 1, \cdots, n_y \\ k = 0, \cdots, \sigma_{U_i}(\lambda_0) + \sigma_{V_j}(\lambda_0) - 1 \end{cases} \tag{25}$$

and

$$\begin{cases} \langle \Phi, X_{\alpha,qt} \rangle = \langle H, X_{\alpha,qt} \rangle \\ \langle \Phi, X_{\beta,pt} \rangle = \langle H, X_{\beta,pt} \rangle \end{cases} \quad \text{for} \begin{cases} i = n_u + 1, \cdots, n_z \\ j = n_y + 1, \cdots, n_w \\ q = 1, \cdots, n_w \\ p = 1, \cdots, n_z \\ t = 0, 1, 2, \cdots \end{cases}. \tag{26}$$

Each of these equations can be viewed as a linear equality constraint on the sequence Φ.

At this point it is convenient to drop the tensor notation used so far and introduce a more compact, computer-ready matrix notation. Let M_{ij} denote an infinite matrix mapping l_1 to $\mathbb{R}^{c_z}$, formed by collecting those coefficients of the zero interpolation functionals that act on the sequence ϕ_{ij}. Similarly, define $\overline{M}_{ij}$ to be an infinite matrix mapping l_1 to l_1, formed by collecting those coefficients of the rank interpolation functionals that act on ϕ_{ij}. With this notation, the set of feasible closed-loop maps is characterized by the following set of equality constraints:

$$\sum_{i=1}^{n_z} \sum_{j=1}^{n_w} M_{ij} \phi_{ij} = \sum_{i=1}^{n_z} \sum_{j=1}^{n_w} M_{ij} h_{ij} =: b_1 \in \mathbb{R}^{c_z} \tag{27}$$

$$\sum_{i=1}^{n_z} \sum_{j=1}^{n_w} \overline{M}_{ij} \phi_{ij} = \sum_{i=1}^{n_z} \sum_{j=1}^{n_w} \overline{M}_{ij} h_{ij} =: b_2 \in l_1. \tag{28}$$

Therefore, the primal optimization problem (5) is equivalent to the following infinite dimensional linear program:

$$\mu^o := \min_{\mu, \xi, \phi_{ij}^+, \phi_{ij}^-} \mu$$

subject to

$$\xi(i) + \sum_{j=1}^{n_w} \sum_{t=0}^{\infty} \phi_{ij}^{-}(t) + \phi_{ij}^{-}(t) = \mu \qquad \text{for } i = 1,\cdots,n_z$$

$$\sum_{i=1}^{n_z} \sum_{j=1}^{n_w} M_{ij}(\phi_{ij}^{+} - \phi_{ij}^{-}) = b_1$$

$$\sum_{i=1}^{n_z} \sum_{j=1}^{n_w} \overline{M}_{ij}(\phi_{ij}^{+} - \phi_{ij}^{-}) = b_2$$

$$\xi, \phi_{ij}^{+}, \phi_{ij}^{-} \geq 0 \tag{29}$$

where $\xi \in \mathbb{R}^{n_z}$ is a positive vector of slack variables. Note that the above linear program is infinite dimensional in the number of variables (i.e., dimension of any ϕ_{ij}) and the number of constraints (i.e., dimension of b_2).

In order to complete this discussion, it remains to show that problem (7) is also equivalent to a linear programming problem. In fact, it can be shown that such problem corresponds to the standard dual formulation of problem (29). To illustrate this fact, we will simply write the dual form of (29) and compate it to (7). Let $\gamma \in l_\infty$ denote the sequence of dual variables. To get more insight into the dual problem, let us partition γ according to the natural partitioning of the set of equality constraints. That is, let $\gamma =: (-\gamma_0 \ \gamma_1 \ \gamma_2)^T$, where $\gamma_0 \in \mathbb{R}^{n_z}$, $\gamma_1 \in \mathbb{R}^{c_z}$ and $\gamma_2 \in l_\infty$ (it is convenient to have the sign of γ_0 changed). Then, the standard dual linear program of (29) is given by:

$$\mu^o = \max_{\gamma_0, \gamma_1, \gamma_2} \langle b_1, \gamma_1 \rangle + \langle b_2, \gamma_2 \rangle$$

subject to

$$\gamma_0 \geq 0, \quad \sum_{i=1}^{n_z} \gamma_0(i) \leq 1$$

$$-\gamma_0(i) \leq \left(M_{ij}^T \gamma_1 + \overline{M}_{ij}^T \gamma_2\right)(k) \leq \gamma_0(i)$$

$$\text{for} \begin{cases} i = 1,\cdots,n_z \\ j = 1,\cdots,n_w. \\ k = 0,1,\cdots \end{cases} \tag{30}$$

If one compares the above linear program with problem (7), the following relationships become apparent: 1) γ_1 and γ_2 are nothing but the coefficients that combine the linear functionals associated with the zero interpolation conditions and the rank interpolation conditions, respectively, to obtain G; 2) the objective function results from expanding $\langle H, G \rangle$ when G is expressed as a linear combination of the elements in the generator of $\mathcal{S}^{\perp}$ with coefficients (γ_1, γ_2); and 3) the set of inequality constraints is equivalent to $\|G\|_\infty \leq 1$, while the second line of inequalities bounds G componentwise, the first line bounds the matrix ∞-norm of G by one.

VI. One-Block Problems

One-block problems have a very specific interpolation structure, namely no rank interpolation conditions. From a primal formulation point of view [see (29)], this simplifies the problem significantly by bringing the number of equality constraints down to a finite value, namely $c_z + n_z$. There remains, however, an infinite number of variables represented by the ϕ_{ij}'s in l_1. Nevertheless, it has been shown by looking at the structure of the dual problem, that the underlying problem is finite dimensional [9]. Indeed, the dual formulation has an infinite number of inequality constraints but retains a finite number of variables:

$$\mu^o = \max_{\gamma_0, \gamma_1} \langle b_1, \gamma_1 \rangle$$

subject to

$$\gamma_0 \geq 0; \quad \sum_{i=1}^{n_z} \gamma_0(i) \leq 1$$

$$-\gamma_0(i) \leq \left(M_{ij}^T \gamma_1\right)(k) \leq \gamma_0(i) \qquad \text{for} \begin{cases} i = 1,\cdots,n_z \\ j = 1,\cdots,n_w. \\ k = 0,1,\cdots \end{cases} \tag{31}$$

Recall that M_{ij}^T is the matrix representation of an operator mapping $\mathbb{R}^{c_z}$ to l_∞. However, with Assumption 1 holding, the actual range of M_{ij}^T is in c_0 since each of the columns of M_{ij}^T is in c_0 and there are only finitely many of them. This is exploited in the following lemma from [34].

Lemma 6.1: Let M be a full column rank infinite matrix mapping $\mathbb{R}^n$ to c_0. Then there exists a positive integer N such that

$$\|(I - P_N)Mx\|_\infty < \|P_N Mx\|_\infty$$

for all nonzero $x \in \mathbb{R}^n$.

Note, in particular, that N is independent of x and is only a function of M.

In other words, given a matrix mapping a finite dimensional space to c_0, it is always possible to bound the index at which the infinity norm of any sequence in the range is achieved.

The following theorem extends a result from [9] by exploiting this structure.

Theorem 6.1: The exact solution of a one-block l_1-optimal control problem is given by the following finite dimensional (dual) linear program,

$$\mu^o = \max_{\gamma_0, \gamma_1} \langle b_1, \gamma_1 \rangle$$

subject to

$$\gamma_0 \geq 0, \quad \sum_{i=1}^{n_z} \gamma_0(i) \leq 1$$

$$-\gamma_0(i) \leq \left(M_{ij}^T \gamma_1\right)(k) \leq \gamma_0(i)$$

$$\text{for} \begin{cases} i = 1,\cdots,n_z \\ j = 1,\cdots,n_w \\ k = 0,\cdots,N_{ij} < \infty \end{cases}. \tag{32}$$

Proof: Form matrices M_{ij}^T as defined before. Assume they have full column rank (if not reduce the number of columns). Apply Lemma 6.1 to each M_{ij}^T and let N_{ij} denote the corresponding index bound. Then, we claim that for every feasible solution of problem (31) all inequalities of the form $|(M_{ij}^T\gamma_1)(k)| \leq \gamma_0(i)$ for $k > N_{ij}$ are inactive constraints (i.e., the inequality is strict) and they can be ignored in the solution. Indeed, by Lemma 6.1, if there is an active constraint for $k > N_{ij}$, then there must have been a violation of a constraint for some $k < N_{ij}$ since the l_∞ norm of the sequence $M_{ij}^T\gamma_1$ is attained before N_{ij} and is always bounded by $\gamma_0(i)$. ■

This fact has an immediate and important implication on the primal linear programming formulation of one-block problems. Due to the alignment conditions, if a dual optimal solution is such that all inequality constraints are inactive for $k > N$, then the primal optimal solution is such that it vanishes for $k > N$.

Corollary 6.1: For any one-block problem, the l_1-optimal closed-loop response, Φ^o, has finite support (i.e., finite pulse response). Furthermore, each entry ϕ_{ij} has support no greater than N_{ij}.

Note that the N_{ij}'s provide apriori bounds on the lengths of the optimal ϕ_{ij}'s. Moreover, these bounds are independent of H and only depend on the zero interpolation structure of the problem.

We conclude this section with an interesting property of most one-block problems, regarding the l_1-norm of each row of the optimal solution.

Corollary 6.2: Given a one-block problem, if for some $i \in \{1,\cdots,n_z\}$ and $j \in \{1,\cdots,n_w\}$ the matrix M_{ij}^T has full column rank, then $\|(\Phi^o)_i\|_1 = \mu^o$.

Proof: Assume $\|(\Phi^o)_i\|_1 < \mu^o$, then $\xi(i) > 0$. By the alignment conditions, this implies that $\gamma_0(i) = 0$, and in view of (32) and the rank condition on M_{ij}^T, γ_1 must be zero. But this implies that $\mu^o = 0$ which is a contradiction. ■

It should be noted that there are some pathological cases where the rank condition of M_{ij}^T is violated. For instance, if the given one-block problem is in fact a combination of two or more totally decoupled subproblems, then some M_{ij}^T's will have entire columns of zeros. In most cases, however, the solution is such that the norm of each row of Φ^o is equal to μ^o. It is interesting to point out the analogy between this aspect of the l_1-optimal solution of one-block problems, and the equivalent in $\mathscr{H}_\infty$ optimization. In the first one, the same "gain" is achieved at all outputs while in the second one the same "gain" is achieved at all frequencies (i.e., inner solution). These are direct consequences of the corresponding norm definitions. Furthermore, the analogy extends to the multiblock case in the sense that this property does not hold in general.

VII. Multiblock Problems

The exact solution of the one-block problem rests on the fact that the primal linear programming formulation has only finitely many equality constraints (or, equivalently, the dual formation has finitely many variables). The multiblock problem, however, is characterized by a primal and dual formulation with an infinite number of variables and constraints. So, in principle, one can attempt to get approximate solutions by an appropriate truncation of the original problem.

There are basically two approximation methods reported in the literature. The first one, known as the *finitely many variables* (FMV) approximation, was originally introduced in [11] and further developed in [29], [34]. It results from constraining the support of the closed-loop response Φ, thus providing a suboptimal finitely supported feasible solution to the problem. In the second approach, known as the *finitely many equations* (FME) approximation [6], [33], only finitely many equality constraints are retained in the primal formulation of the problem, the solution of which is superoptimal but infeasible. Its value is complementary to the first approach in the sense that it generates lower bounds of the optimal norm, μ^o.

The next two subsections give a more detailed description of these methods along with their main characteristics. They do not contain new results.

A. The FMV Approximation Method

Let N be the order of approximation or support of Φ, then the FMV primal formulation is given by the following linear program:

$$\overline{\nu}_N := \min_{\mu,\,\xi,\,\phi_{ij}^+,\,\phi_{ij}^-} \mu$$

subject to

$$\xi(i) + \sum_{j=1}^{n_w}\sum_{k=0}^{N} \phi_{ij}^-(k) + \phi_{ij}^-(k) = \mu \qquad \text{for } i = 1,\cdots,n_z$$

$$\sum_{i=1}^{n_z}\sum_{j=1}^{n_w} M_{ij}(\phi_{ij}^+ - \phi_{ij}^-) = b_1$$

$$\sum_{i=1}^{n_z}\sum_{j=1}^{n_w} \overline{M}_{ij}(\phi_{ij}^+ - \phi_{ij}^-) = b_2$$

$$\phi_{ij}^+(k) = \phi_{ij}^-(k) = 0 \qquad \text{for } k > N$$

$$\xi,\ \phi_{ij}^+,\ \phi_{ij}^- \geq 0. \tag{33}$$

Note that without the constraints $\phi_{ij}^+(k) = \phi_{ij}^-(k) = 0$ for $k > N$, (33) is equivalent to the full (untruncated) optimization problem. Clearly, the added constraints will make $\overline{\nu}_N \geq \mu^o$ in general. It is yet unclear, however, if the resulting problem is finite dimensional or not, since we still carry an infinite number of constraints. A closer look at the matrices $\overline{M}_{ij}$ will answer this question.

Recall that these matrices represent the rank interpolation conditions (albeit some specific reordering) of the

form (see Theorem 3.1):

$$\begin{pmatrix} \alpha_{n_u+1} \\ \vdots \\ \alpha_{n_z} \end{pmatrix} * \Phi = \begin{pmatrix} \alpha_{n_u+1} \\ \vdots \\ \alpha_{n_z} \end{pmatrix} * H$$

and

$$\Phi * \begin{pmatrix} \beta_{n_y+1} & \cdots & \beta_{n_w} \end{pmatrix} = H * \begin{pmatrix} \beta_{n_y+1} & \cdots & \beta_{n_w} \end{pmatrix}$$

where the results from the right-hand side convolutions are collected in the infinite vector b_2. The matrix representation of the convolution of the α_i's and β_j's on the different entries of Φ, say ϕ_{ij}, is precisely given by $\overline{M}_{ij}$. Therefore, such infinite matrices will have a band structure inherited from the fact that the $\hat{\alpha}_i(\lambda)$'s and $\hat{\beta}_j(\lambda)$'s are polynomials.

In view of this particular structure, forcing $\phi_{ij}(k) = 0$ for $k > N$ will make the product $(\overline{M}_{ij}\phi_{ij})(k)$ eventually vanish for $k > N + constant$, where the constant depends on the order of the polynomials $\hat{\alpha}(\lambda)$'s and $\hat{\beta}(\lambda)$'s. If, however, the infinite vector b_2 is not zero at that point, then the equality constraints will be violated for any Φ, implying that the added constraints have transformed the feasible set into an empty set and that the linear program has no solution. Furthermore, this will always be the case if b_2 has infinite support, no matter how large N is chosen to be. This leads to the following theorem and corollary (equivalent results can be found in [29]).

Theorem 7.1: Given a multiblock problem, there exists a finitely supported feasible solution, Φ, if and only if $\alpha_i * H$ and $H * \beta_j$ are finitely supported for $i = n_u + 1, \cdots, n_z$ and $j = n_y + 1, \cdots, n_w$.

Corollary 7.1: Given a positive integer N, the *FMV* problem (33) has a nonempty feasible set and therefore a solution, if and only if $(\alpha_i * H)(k) = 0$ and $(H * \beta_j)(k) = 0$ for $k > N +$ constant, $i = n_u + 1, \cdots, n_z$ and $j = n_y + 1, \cdots, n_w$, where the constant depends on the order of $\hat{\alpha}_i$ and $\hat{\beta}_j$.

It is clear from the above results that there is a class of multiblock problems for which the FMV method fails regardless of the order of approximation N. Also, given any multiblock problem, there is in general a lower bound for N under which the FMV method also fails. A way to avoid this difficulty is to approximate H arbitrarily close with a finitely supported sequence (e.g., $P_k H$). Such approach, however, has the effect of increasing the order of the suboptimal solution and therefore the order of the controller that achieves it.

Without overlooking these limitations, we are going to assume for the rest of this subsection that the problems at hand allow polynomial feasible solutions and that N is large enough to capture at least one of such solutions.

Under these assumptions, it is clear that all but finitely many constraints in (33) are satisfied trivially, so that the problem is in effect a finite dimensional linear program. The next theorem shows that it has nice convergence properties [11].

Theorem 7.2: In the *FMV* method, $\overline{\nu}_N \to \mu^o$ as $N \to \infty$.

Besides the necessary assumptions regarding the existence of polynomial feasible solutions, the FMV approximation method suffers from two other signigicant drawbacks: 1) Although it provices an upper bound for μ^o and a feasible solution that achieves it, it gives no information about how far away from optimal the solution is, and 2) the compensators obtained with this method suffer from order inflation (i.e., the order of the controller increases with N). These aspects of the solutions will be illustrated through an example at the end of this section.

B. The FME Approximation Method

The first drawback was solved independently in [6] and [33] by introducing a second optimization problem, the FME approximation method. Such method further exploits the structure of the matrices $\overline{M}_{ij}$ to get lower bounds on μ^o. The name stems from the fact that only finitely many equality constraints associated with the rank interpolation conditions are included in the optimization problem. The rest are simply ignored. Therefore, the solution obtained will in general fail to satisfy those constraints that were left out, rendering it infeasible to the un-truncated problem. A formal statement of the FME approximation problem (in its primal form) is as follows:

$$\underline{\nu}_N := \min_{\mu, \xi, \phi_{ij}^+, \phi_{ij}^-} \mu$$

subject to

$$\xi(i) + \sum_{j=1}^{n_w} \sum_{k=0}^{\infty} \phi_{ij}^+(k) + \phi_{ij}^-(k) = \mu \qquad \text{for } i = 1, \cdots, n_z$$

$$\sum_{i=1}^{n_z} \sum_{j=1}^{n_w} M_{ij}(\phi_{ij}^+ - \phi_{ij}^-) = b_1$$

$$\left(\sum_{i=1}^{n_z} \sum_{j=1}^{n_w} \overline{M}_{ij}(\phi_{ij}^+ - \phi_{ij}^-) \right)(k) = b_2(k) \qquad \text{for } k = 0, \cdots, N-1$$

$$\xi, \phi_{ij}^+, \phi_{ij}^- \geq 0 \tag{34}$$

This truncation scheme transforms the original problem into one with a finite number of constraints but still an infinite number of variables. An argument similar to the one used for the one-block problem shows that the above infinite dimensional linear program is indeed equivalent to a finite dimensional one. Let $\overline{M}_{ij,N}$ denote the truncated $\overline{M}_{ij}$ (i.e., the first N rows of it). Since $\overline{M}_{ij,N}$ has only a finite number of rows, then the combined matrix

$$\begin{pmatrix} M_{ij}^T & \overline{M}_{ij,N}^T \end{pmatrix}$$

maps a finite dimensional space to l_∞. Moreover, due to the band structure of $\overline{M}_{ij}$, all the columns of the combined matrix are in c_0 and thus the range is in c_0. Therefore, by Lemma 6.1 and Theorem 6.1, the FME problem is equivalent to a finite dimensional linear program whose solution has finite support.

The sequence of linear programs in (34) are such that the number of constrints increases with N. Therefore, $\underline{\nu}_N$ forms a nondecreasing sequence bounded from above by μ^o. The next theorem shows that it actually converges to μ^o [34].

Theorem 7.3: In the *FME* method, $\underline{\nu}_N \to \mu^o$ as $N \to \infty$.

Based on these convergence properties, a multiblock problem can be solved iteratively to any degree of approximation by solving two finite dimensional linear programs, corresponding to the FMV and FME truncation schemes, at each iteration. The stopping criterion is based on the upper and lower bounds provided in each iteration. This holds only if there exits finitely supported feasible solutions to the problem.

C. Delay Augmentation Method

Following, a new method is presented by the name of delay augmentation (DA). This method provides a conceptually attractive and computationally efficient way of solving general multiblock problems, with the added benefit of not requiring assumptions on the existence of polynomial feasible solutions and with the capacity of generating suboptimal controllers without order inflation.

The main idea is very simple.

1) augment U and V with pure delays (i.e., right shifts) such that the augmented problem is one-block,

2) apply all the machinery developed for one-block problems to the augmented system,

3) reduce it back to the original system and compute the controller.

In more precise terms, partition the original system as follows:

$$\begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} - \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} Q \begin{pmatrix} V_1 & V_2 \end{pmatrix} \quad (35)$$

where $U_1 \in l_1^{n_u \times n_u}$ and $V_1 \in l_1^{n_y \times n_y}$. Then, augment U and V with Nth order shifts and augment the free parameter Q accordingly:

$$\begin{pmatrix} \Phi_{11,N} & \Phi_{12,N} \\ \Phi_{21,N} & \Phi_{22,N} \end{pmatrix} := \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} - \begin{pmatrix} U_1 & 0 \\ U_2 & S_N \end{pmatrix} \cdot \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \begin{pmatrix} V_1 & V_2 \\ 0 & S_N \end{pmatrix} \quad (36)$$

or, equivalently,

$$\Phi_N := H - U_N Q_N V_N =: H - R_N \quad (37)$$

where U_N, Q_N, and V_N have the obvious definitions. Clearly, problem (37) is of the one-block class since $U_N \in l_1^{n_z \times n_z}$ and $V_N \in l_1^{n_w \times n_w}$. By expanding (36) we have

$$\Phi_N = H - UQ_{11}V - S_N \tilde{R}_N \quad (38)$$

and

$$\tilde{R}_N := \begin{pmatrix} 0 & U_1 Q_{12} \\ Q_{21} V_1 & Q_{21} V_2 + U_2 Q_{12} + S_N Q_{22} \end{pmatrix}$$

where the fact that these are all time invariant operators has been used. With this notation we are ready to define the delay augmentation problem of order N as the following optimization problem:

$$\underline{\mu}_N := \inf_{Q_N \in l_1^{n_z \times n_w}} \|H - U_N Q_N V_N\|_1. \quad (39)$$

It follows from the above definition that $\underline{\mu}_N$ is a lower bound for μ^o since

$$\begin{aligned} \underline{\mu}_N &\le \inf_{\substack{Q_{11} \in l_1^{n_u \times n_y} \\ Q_{12} = Q_{21} = Q_{22} = 0}} \|H - U_N Q_N V_N\|_1 \\ &= \inf_{Q_{11} \in l_1^{n_u \times n_y}} \|H - UQ_{11}V\|_1 = \mu^o. \end{aligned}$$

In other words, the extra degree of freedom in the free parameter Q_N (as compated to Q) makes the construction of superoptimal solutions possible. Such solutions, however, are clearly infeasible to the unaugmented problem. Also, it is interesting to note that the extra parameters (namely Q_{12}, Q_{21}, and Q_{22}) have no effect on the solution $\Phi_N(k)$ for $k < N$ due to the presence of the shift operator in (38). And even more interesting, the term Φ_{11} is not affected at all by the added parameters (note the block of zeros in $\tilde{R}_N$). This observation will let us construct a suboptimal feasible solution directly from the solution of (39).

Given some positive integer N, let

$$\underline{\mu}_N = \|\Phi_N^o\|_1 = \|H - UQ_{11}^o V - S_N \tilde{R}_N^o\|_1$$

then, clearly

$$\mu^o = \inf_{Q \in l_1^{n_u \times n_y}} \|H - UQV\|_1 \le \|H - UQ_{11}^o V\|_1 =: \overline{\mu}_N. \quad (40)$$

Or, equivalently, the solution obtained by making the extra free parameters zero after solving (39) is feasible and suboptimal to the unaugmented problem. The following lemma summarizes these results.

Lemma 7.1: Given a positive integer N and definitions (39) and (40), then

$$\underline{\mu}_N \le \mu^o \le \overline{\mu}_N$$

where $\overline{\mu}_N$ is achieved with Q_{11}^o.

Before addressing the convergence properties of this method, a word on existence is in order. Recall that existence is assured if there are no zero interpolations on the boundary of the unit disk. Now, it may happen that a multiblock problem that satisfies this condition augments into a one-block problem that does not. Indeed, notice that the left zeros of $\hat{U}_N$ are given by the left zeros of $\hat{U}_1$ plus a multiple zero at the origin (due to the block of delays, $\lambda^N I$, resulting from the λ-transform of S_N). Clearly, the left zeros of $\hat{U}$ are also left zeros of $\hat{U}_1$. However, $\hat{U}_1$ may have more zeros, possibly on the boundary of the

disk. For example, let

$$\hat{U}_N = \begin{pmatrix} (\lambda - 1) & 0 \\ (\lambda - 0.5) & \lambda^N \end{pmatrix}.$$

At $\lambda = 1$ the above matrix looses rank, indicating the existence of a zero at the boundary of the unit disk. However, reordering the outputs before augmenting with delays avoids this difficulty:

$$\hat{U}_N = \begin{pmatrix} (\lambda - 0.5) & 0 \\ (\lambda - 1) & \lambda^N \end{pmatrix}.$$

Note that the original $\hat{U}$ has no left zeros since the rows are coprime.

The same applies to the right zeros of $\hat{V}$. In many instances this situation may be reversed by a proper reordering of inputs and outputs, such that the resulting $\hat{U}_1$ and $\hat{V}_1$ have no zeros on the boundary respectively. In any case, this limitation has little practical implications since it is always possible to find rational solutions to (39) that are arbitrarily close to $\underline{\mu}_N$. In view of this, we will make the following simplifying assumption.

Assumption 2: $\hat{U}_1(\lambda)$ and $\hat{V}_1(\lambda)$ have no zeros on the unit circle.

Note that under this assumption the results of Theorem 3.1 are applicable. Furthermore, in the analysis that follows we will be able to exploit the existence of optimal solutions for any N and thus avoid the epsilon-delta arguments that would result from rational approximations.

By definition, problem (39) is equivalent to the following primal-dual pair:

$$\underline{\mu}_N = \min_{R_N \in \mathcal{S}_N} \|H - R_N\|_1 = \sup_{\substack{G_N \in \mathcal{S}_N^\perp \\ \|G_N\|_\infty \le 1}} \langle H, G_N \rangle. \quad (41)$$

It is easy to see that, as N increases, the subspace $\mathcal{S}_N$ gets smaller and such that

$$\mathcal{S}_N \supseteq \mathcal{S}_{N+1} \supseteq \cdots \supseteq \mathcal{S} \quad (42)$$

since the only change in the interpolation structure is due to a higher multiplicity of the zero at the origin. Therefore, $\underline{\mu}_N$ forms a nondecreasing sequence, bounded from above by μ^o.

The next theorem states an interesting convergence result.

Theorem 7.4: Given the sequence Φ_N^o, there exists a subsequence that converges weak* to some Φ^o. If the optimal solution is unique then the whole sequence converges weak* to it.

Proof: Clearly Φ_N^o forms a bounded sequence in $l_1^{n_z \times n_w}$, then there exists a weak*-convergent subsequence $\Phi_{N_s}^o$, by the Banach–Alaoglu theorem. Let Φ^{w^*} denote such limit point. As mentioned before, Φ_N^o is infeasible to the original (unaugmented) problem. However, we will show that Φ^{w^*} is in fact feasible. From (38), after taking the weak* limit, we have:

$$\Phi^{w^*} = H(UQ_{11}^o V)^{w^*} - \left(S_{N_s}\tilde{R}_{N_s}^o\right)^{w^*} = H - U(Q_{11}^o)^{w^*} V$$

where the superscript w^* denotes *weak** limit. The last term drops since $\tilde{R}_N^o$ is uniformly bounded in N. For if $\{\tilde{R}_N^o\}$ were unbounded, then $\{Q_{11}^o\}$ would necessarily be unbounded to keep $\underline{\mu}_N$ bounded. But this contradicts the fact that $\underline{\mu}_N$ is larger than $\|H_{11} - U_1 Q_{11,N}^o V_1\|_1$. Therefore, Φ^{w^*} is feasible. To show that Φ^{w^*} is actually an optimal solution, we need to view Φ_N^o as a bounded linear operator from $c_0^{n_z \times n_w}$ to $\mathbb{R}$ (i.e., bounded linear functional on $c_0^{n_z \times n_w}$) with strong operator limit Φ^{w^*}. In such context we have the following inequality (see [25], p. 269):

$$\|\Phi^{w^*}\|_1 \le \liminf_{s \to \infty} \|\Phi_{N_s}^o\|_1 \le \|\Phi^o\|_1.$$

Therefore, since Φ^{w^*} is feasible, all inequalities above are in fact equalities and $\Phi^{w^*} = \Phi^o$.

Finally, if the solution is unique then the whole sequence converges to Φ^o *weak**. ∎

The last claim in the above lemma simply reflects the fact that if there are several optimal solutions, Φ^o, then a sequence of DA problems can be such that Φ_N^o (in the limit) "jumps" from one optimal solution to the other therefore not converging as a whole. Then, a subsequence that "keeps track" of a single optimal solution will converge *weak** to it. This technicality is unnecessary when the optimal solution in unique.

An immediate corollary to Theorem 7.4 is the following.

Corollary 7.2: The sequence of lower bounds, $\underline{\mu}_N$, converge to μ^o as $N \to \infty$.

Next, we focus on the convergence properties of the dual sequence G_N. In the context of (41) we state the following Theorem. (Note that G_N^o as well as G^o may not be unique).

Theorem 7.5: Given the sequence G_N^o, there exists a subsequence that converges *weak** in $l_\infty^{n_z \times n_w}$ to an optimal solution G^o. Furthermore, if the solution G^o is unique, then the whole sequence converges *weak** to it.

Proof: Clearly the sequence G_N^o is bounded by one. Then, by the Banach–Alaoglu theorem, there exists a subsequence that converges *weak** in $l_\infty^{n_z \times n_w}$. Also, from (42) we have that

$$\mathcal{S}_N^\perp \subseteq \mathcal{S}_{N+1}^\perp \subseteq \cdots \subseteq \mathcal{S}^\perp.$$

Or, equivalently, G_N^o is feasible to the original (dual) problem for all N. Further, it can be shown that the feasible subspace $\mathcal{S}^\perp$ is *weak**-closed [11], [29], then $G_{N_s}^o$ converges *weak** to a feasible limit point, say G^{w^*}. Therefore,

$$\underline{\mu}_{N_s} = \langle H, G_{N_s}^o \rangle \to \langle H, G^{w^*} \rangle.$$

But, by Corollary 7.2, $\underline{\mu}_{N_s} \to \mu^o$, thus, $\mu^o = \langle H, G^{w^*} \rangle$. This implies that G^{w^*} is in fact an optimal dual solution, G^o, since it achieves the optimal value and is feasible.

If the solution, G^o, is unique then the whole sequence converges *weak** to it. ∎

Next, we focus our attention on the sequence of suboptimal solutions that attain the upper bound $\overline{\mu}_N$. Let

$\overline{\Phi}_N := H - UQ_{11}^o V$, then $\overline{\mu}_N = \|\overline{\Phi}_N\|_1$ by definition. It is easy to see that $\overline{\Phi}_N$ forms a bounded sequence in $l_1^{n_z \times n_w}$ (if not $\Phi_{11,N}^o$ and thus $\underline{\mu}_N$ would be unbounded). Therefore, there exists a subsequence that converges *weak** in $l_1^{n_z \times n_w}$. Also, $\overline{\Phi}_N$ is clearly feasible to the original problem for any N, and since $\mathcal{S}$ is *weak**-closed [11], then all *weak** limit points are feasible. The question is whether or not the subsequence $\overline{\mu}_{N_s} = \|\overline{\Phi}_{N_s}\|_1$ converges to μ^0 in general.

In order to give a proper answer to this question, it is useful to make the following observation first made in [33]. In Corollary 6.2 we have shown that most one-block problems have optimal solutions with all row norms equal to μ^o. To illustrate why this is not the case with multiblock problems, consider the following SISO example:

$$\phi_1 = h_1 - u_1 q$$

where all operators are in l_1 and $\hat{u}_1(\lambda)$ has no zeros on the unit circle. Let ϕ_1^o denote an l_1-optimal solution to such (one-block) problem, that is achieved with q^o. Next, add a new row to the problem,

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} - \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} q$$

such that $\|h_2 - u_2 q^o\|_1 < \mu^o$ (this is always possible simply by choosing a scalar weight on the second row of small enough value). Then, it is clear that an optimal solution to the new two-block column problem is still given by q^o and that $\|\phi_2^o\|_1 < \|\phi_1^o\|_1 = \mu^o$. In other words, the new row does not affect the optimal solution which is given by the first row alone. In contrast with a one-block problem with two outputs, a two-block problem with two outputs has to minimize both outputs with just one scalar free parameter sequence, q. The "shortage" of degrees of freedom is what makes this situation more common in multiblock problems.

Having noted this behavior, we can present the main theorem concerning convergence of the upper bound, $\overline{\mu}_N$.

Theorem 7.6: Given a general multiblock problem, let $\Phi_{N_s}^o$ converge *weak** to an optimal solution $\Phi^o = H - UQ^oV$ such that $\|(\Phi^o)_i\|_1 = \mu^o$ for $i \in \{1, \cdots, n_u\}$. Then, $\overline{\Phi}_{N_s}$ converges strongly (i.e., in the norm) to Φ^o as $N \to \infty$, and further, $\overline{\mu}_N \to \mu^o$.

Proof: It is a well-known fact that if a sequence $x_n \in l_1$ converges to x^{w^*} *weak**, and if $\|x_n\|_1 \to \|x^{w^*}\|_1$, then x_n converges to x^{w^*} strongly. However, such result is valid only for scalar and row-vector sequences in l_1 (it is easy to think of a counter-example in the general matrix case). Therefore, we apply it to each individual row of Φ_{N_s} to conclude the following: $(\Phi_{N_s}^o)_i$ converges strongly (i.e., in the norm) to $(\Phi^o)_i$ for all $i \in \{1, \cdots, n_z\}$ such that $\|(\Phi^o)_i\|_1 = \mu^o$.

At the same time, from Assumption 2, $\hat{U}_1$ and $\hat{V}_1$ have full normal rank, so the map from Q_{11} to $\Phi_{11,N}$ is continuous with continuous inverse, that is

$$\hat{Q}_{11} = \hat{U}_1^{-1}\left(\hat{H}_{11} - \hat{\Phi}_{11,N}\right)\hat{V}_1^{-1}.$$

Then, using the fact that $\|(\Phi^o)_i\|_1 = \mu^o$ for $i \in \{1, \cdots, n_u\}$, we conclude that Φ_{11,N_s}^o converges strongly to Φ_{11}^o which in turn implies that Q_{11}^o converges strongly to Q^o and the result follows. ∎

The above theorem suggests that the construction of the feasible solution that attains the upper bound, $\overline{\Phi}_N$, can be viewed as an attempt to compute the *weak** limit of the sequence $\Phi_{N_s}^o$ by "throwing away the tail" contained in the term $S_N \tilde{R}_N^o$.

It should be stressed at this point, that nonpathological multiblock problems have optimal solutions where at least n_u of the n_z rows achieve the optimal norm (a natural extension of how optimal solutions of one-block problems behave). Furthermore, those rows that do not achieve the optimal norm can be left out of the optimization problem without affecting the overall solution, so eventually, the problem can be reduced. In general, however, a well posed control problem will tend to have none of its rows "redundant", so $\overline{\mu}_N$ usually converges to μ^o without further considerations. In this context we have the following corollary valid for two-block column problems of the form:

$$\begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix} - \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} QV.$$

Corollary 7.3: Given a two-block column problem, if $\|\Phi_2^o\|_1 < \mu^o$ then $\overline{\Phi}_N$ is the exact optimal solution for any N.

Proof: Follows immediately from the fact that the first block-row $H_1 - U_1 Q_1 V$ is independent of the extra free parameter. That is,

$$\Phi_{1,N}^o = H_1 - U_1 Q_1^o V$$
$$\Phi_{2,N}^o = H_2 - U_2 Q_1^o V - S_N Q_2^o V.$$

Then, for any N we have

$$\|\Phi_{1,N}^o\|_1 \geq \mu^o \geq \underline{\mu}_N = \max\left(\|\Phi_{1,N}^o\|_1, \|\Phi_{2,N}^o\|_1\right) \geq \|\Phi_{1,N}^o\|_1.$$

Thus, equality is attained throughout and the result follows, i.e., $Q_1^o = Q^o$. ∎

Theorem 7.6 and Corollary 7.3 dictate that a reordering of outputs needs to be done so that the first n_u rows of Φ achieve the optimal norm μ^o. The question is, then, how to find *a priori* which rows of the problem are not going to achieve the optimal norm. A brute force answer to this question is simply to solve all possible one-block problems formed by taking n_u rows out of the given n_z rows. If any solution is such that all the rows that were left out have smaller norm than the corresponding μ^o, then those rows are the inactive ones and should be ordered in U_2. (In fact these rows can be removed altogether.) However, this approach may require a considerable amount of work. We will return to this difficulty later.

Two-block row problems, show a similar behavior. Indeed, such problems may have columns that are inactive in the optimization process in the sense that they can be

removed without affecting the solution. Note that in the previous case, the phenomenon of inactive rows was intimately related to the fact that the l_1 norm on matrices takes the maximum row norm, which allowed us to easily construct an example.

If the DA method is applied to a two-block row problem such that the columns associated with V_2 are inactive, then again the solution $\overline{\Phi}_N$ is exact for any N. However, $\underline{\mu}_N$ will not give the exact optimal norm (although it will tend to it) since the extra parameter contributes in reducing the norm of $\Phi^o_{12,N}$.

Finally, let us point out that both forms of redundancy (row and column) can occur in a multiblock problem simultaneously. This discussion motivates the following definitions.

Definition 7.1: Given a general multiblock problem, a one-block partition is defined by taking n_y inputs and n_u outputs of the full problem, such that the reduced problem corresponds to a one-block problem with full normal rank U and V.

Definition 7.2: In a multiblock problem, a one-block partition is totally dominant (TD) if the optimal free parameter Q^o obtained from its solution also solves the original multiblock optimization problem.

It follows from these definitions that, if there is a TD one-block partition corresponding to the partitions U_1 and V_1, then the DA method provides the exact answer for any N. The next section illustrates some of these properties.

In summary, in the DA method, $\underline{\mu}_N$ always converges to μ^o, and $\overline{\mu}_N$ converges to μ^o if the first n_u rows are active.

VIII. A Comparison of Methods

This section provides a general comparison of the approximation methods presented, based on a few simple multiblock examples. To facilitate further study, the first two selected problems are the same as those treated in other references [11], [33]. Particular attention will be paid to two aspects of the solutions: first, the support characteristics of the sequence of solutions, and second, the order of the suboptimal controller they generate.

Example I: Consider the following two-block column problem: given the SISO plant P, minimize the l_1 norm of the weighted sensitivity and complementary sensitivity,

$$\Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} W_1(1-PK)^{-1} \\ W_2PK(1-PK)^{-1} \end{pmatrix}$$

where

$$\hat{P}(\lambda) = \frac{\lambda(\lambda - 0.5)}{(\lambda - 0.1)(1 - 0.5\lambda)}$$

and

$$\hat{W}_1(\lambda) = \frac{0.02}{1 - 0.2\lambda}; \quad \hat{W}_2(\lambda) = \frac{0.004\rho}{1 - 0.6\lambda}.$$

Note that a variable scalar weight on ϕ_2, denoted ρ, has been included. By adjusting ρ, we will be able to generate two interesting cases: case a) where ϕ_1 is TD (for "small" ρ) and case b) where both rows are active in the optimization (for "intermediate" ρ). The workings of Theorem 7.6 will be illustrated by reordering the outputs and forcing the TD row to be in the "wrong" place.

The results are presented in tables showing, for each N, the DA lower bound ($\underline{\mu}_N$), the DA upper bound ($\overline{\mu}_N$) and the FMV upper bound ($\overline{\nu}_N$). The FME lower bound is omitted since it is equal to $\underline{\mu}_N$ in this particular case. In general, however, $\underline{\mu}_N$ converges faster than $\underline{\nu}_N$ since the delay augmentation method generates more constraints than the FME method for any given N. These extra constraints are the ones that ensure feasibility of $\Phi_{11,N}$. To illustrate this point, consider the following case:

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} - \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} q$$

where $\hat{u}_1(\lambda)$ and $\hat{u}_2(\lambda)$ are coprime. Further, assume that $\hat{u}_1(\lambda)$ has an unstable zero at λ_0. Consequently, the FME method generates the following rank constraints (note there are no left zeros of $\hat{U}$):

$$(\phi_1 * u_2 - \phi_2 * u_1)(k) = (h_1 * u_2 - h_2 * u_1)(k); \quad k = 0, \cdots, N-1. \quad (43)$$

Now consider the DA method of order N:

$$\hat{U}_N = \begin{pmatrix} \hat{u}_1 & 0 \\ \hat{u}_2 & \lambda^N \end{pmatrix}.$$

Let us construct the left zero interpolations for this $\hat{U}_N$. Multiplying $\hat{U}_N$ on the left by $(\hat{u}_1 \ -\hat{u}_1)$ we get $(0 \ -\hat{u}_1\lambda^N)$ This implies that the left zeros of $\hat{U}_N$ are given by the zeros of $\hat{u}_1$ and a zero at the origin of multiplicity N. Further, the directional properties of such zero are captured by the vector $(\hat{u}_2 \ -\hat{u}_1)$. Therefore, the zero interpolation conditions are given by (43) plus the following:

$$\hat{\phi}_1(\lambda_0) = \hat{h}_1(\lambda_0).$$

Note that this last constraint becomes redundant as $N \to \infty$.

In this particular numerical example, however, both lower bounds are equal due to the fact that the unstable zeros of $\hat{u}_1(\lambda)$ are also zeros of $\hat{u}_2(\lambda)$.

Also included are the support characteristics of Φ^o_N and of the FMV solution along with the order of the suboptimal controllers that achieve the corresponding upper bounds.

To describe the support characteristics we define a function, len$(\cdot)$, mapping $l_1^{n \times m}$ to $Z_+^{n \times m}$ in the following way: given $\Phi \in l_1^{n \times m}$, then $[\text{len}(\Phi)]_{ij}$ is a nonnegative integer equal to the maximum k for which $\phi_{ij}(k)$ is not zero, plus one. Also, we denote the order of a controller K by ord(K).

Case a): In this case let $\rho = 1$ and keep the same ordering of outputs as above (i.e., sensitivity first). The results are shown in Table I. Clearly, the solution given by the delay augmentation method is exact since the upper

TABLE I
COMPARISON OF METHODS: EXAMPLE I. CASE a) WHERE THE FIRST ROW IS TD

	DA				FMV		
N	$\underline{\mu}_N$	$\overline{\mu}_N$	$\text{len}(\Phi_N^o)^T$	$\text{ord}(K)$	$\overline{\nu}_N$	$\text{len}(\Phi_N^o)^T$	$\text{ord}(K)$
1	0.78222	0.78222	(3 2)	2	—	—	—
2	0.78222	0.78222	(3 3)	2	—	—	—
3	0.78222	0.78222	(3 4)	2	1.31912	(4 4)	4
4	0.78222	0.78222	(3 5)	2	0.97459	(5 5)	5
5	0.78222	0.78222	(3 6)	2	0.87547	(6 6)	6
6	0.78222	0.78222	(3 7)	2	0.83292	(7 7)	7

and lower bounds are equal for any N. Then, in the context of Corollary 7.3, the first row corresponding to the weighted sensitivity is TD. Indeed, a simple computation shows that $\|\phi_2^o\|_1 = 0.2040 < \|\phi_1^o\| = 0.7822$. Note how the support of the second row of the augmented optimal solution increases with N while the first row remains constant and equal to the optimal of the un-augmented problem. Since the controller is computed from the first row only, it is also exact and constant as N increases. In contrast, the FMV solution has increasing support on both rows, thus generating a suboptimal controller of increasing order that approximates the second order optimal controller. Note that for some N's, the FMV problem has no solution (indicated with a dash) since the feasible set is empty.

Next, consider the same problem but with the outputs reordered (i.e., the complementary sensitivity in the first row). Table II shows how violating the conditions of Theorem 7.6 affects the convergence of the upper bound (note that the lower bound does converge as shown in Theorem 7.4). Although the upper bound does not converge, it is interesting to note that for $N \geq 2$ the length of $\phi_{2,N}^o$ (i.e., the weighted sensitivity) locks at a value of 3, which coincides with the length of the optimal solution. This seems to be a general characteristic of the DA method as we shall see later. At the same time, there is a clear order inflation on the suboptimal controller due to the constant increase in the length of $\phi_{1,N}^o$. (Note: FMV results are not included in Table II since such method is not affected by reordering.)

Case b: Let $\rho = 6$ and place the sensitivity back in the first row. For this weighting, both rows are active in the optimization as shown by the gradual convergence of the upper and lower bound (see Table III). Note that, even though the controller order growth is comparable in both methods, the support characteristics are quite different. Most interesting, the length of $\phi_{2,N}^o$ remains equal to 4 for $N > 2$ suggesting the possibility that, by changing the order of the outputs, a low order suboptimal controller can be computed. This is in fact the case, as shown in Table IV. (This procedure does not apply to the FMV method since the suboptimal solutions obtained by this method are such that all entries of $\Phi(k)$ are supported at $k = N$.) It is interesting how in both cases a) and b), a proper ordering of the outputs results in a much better approximation of the solution (exact if one row is TD) in the sense that, after some N, the sequence of suboptimal

TABLE II
COMPARISON OF METHODS: EXAMPLE I, CASE a) WHERE THE SECOND ROW IS TD

N	$\underline{\mu}_N$	$\overline{\mu}_N$	$\text{len}(\Phi_N^o)^T$	$\text{ord}(K)$
1	0.22000	1.1602	(3 2)	2
2	0.29195	1.9939	(4 3)	4
3	0.42826	3.1464	(5 3)	5
4	0.55995	3.9859	(6 3)	6
5	0.65664	4.5189	(7 3)	7
6	0.71550	4.8077	(8 3)	8
7	0.74789	4.9504	(9 3)	9
8	0.76483	5.0171	(10 3)	10
⋮	⋮	⋮	⋮	⋮
15	0.78159	5.1878	(15 3)	15

TABLE III
COMPARISON OF METHODS: EXAMPLE I, CASE b) WHERE NO ROW IS TD

	DA				FMV		
N	$\underline{\mu}_N$	$\overline{\mu}_N$	$\text{len}(\Phi_N^o)^T$	$\text{ord}(K)$	$\overline{\nu}_N$	$\text{len}(\Phi_N^o)^T$	$\text{ord}(K)$
1	0.78222	1.2243	(3 2)	2	—	—	—
2	0.79333	1.2547	(4 3)	3	—	—	—
3	0.90230	1.5255	(4 4)	5	1.3191	(4 4)	3
4	0.99522	1.0389	(5 4)	5	1.0564	(5 5)	4
5	1.0015	1.0105	(6 4)	6	1.0121	(6 6)	6
6	1.0024	1.0043	(7 4)	7	1.0044	(7 7)	7
7	1.0026	1.0030	(8 4)	8	1.0030	(8 8)	8
8	1.0026	1.0027	(9 4)	9	1.0027	(9 9)	9

TABLE IV
COMPARISON OF METHODS: EXAMPLE I, CASE b) WITH THE OUTPUTS REORDERED

N	$\underline{\mu}_N$	$\overline{\mu}_N$	$\text{len}(\Phi_N^o)^T$	$\text{ord}(K)$
1	0.95745	1.1602	(3 2)	2
2	0.95745	1.1602	(3 3)	2
3	0.98658	1.0586	(4 4)	3
4	0.99889	1.0157	(4 5)	3
5	1.0019	1.0053	(4 6)	3
6	1.0022	1.0031	(4 7)	3
7	1.0026	1.0027	(4 8)	3
8	1.0026	1.0026	(4 9)	3

controllers are of fixed order and asymptotically approaching the optimal one. This is not an isolated case. Many other multiblock problems for which reliable numerical approximations were computed behave in this way when solved by the DA method. In other words, given a general multiblock problem, there seems to be a one-block partition that preserves a polynomial optimal solution, and further, such support structure is eventually captured by the delay augmentation method for a large enough N. Then, a proper ordering of inputs and outputs that places the one-block partition in the first n_u rows and n_y columns of Φ (corresponding to U_1 and V_1) will generate a sequence of suboptimal controllers without order inflation.

These observations suggest that an efficient algorithm for computing low order suboptimal controllers can be as follows: given a general multiblock problem,

Step 1) Pick a positive integer N.

Step 2) Solve the corresponding delay augmentation problem.

Step 3) Compute $\operatorname{len}(\Phi_N^o)$ and reorder inputs such that the set of $n_u \times n_y$ input–output pairs of minimum length correspond to Φ_{11}.

Step 4) If reordering was necessary in Step 3), solve the reordered system for the same N. Then, check the difference between the upper and lower bounds, i.e., $\overline{\mu}_N - \underline{\mu}_N$. If such difference is small enough stop, otherwise increase N by one (or more) and go to Step 2).

In order to illustrate the workings of such algorithm we include a four-block example.

Example II: Consider the following 2-input-2-output four-block problem where the regulated signals are the output of the plant and the control sequence (weighted with the scalar ρ), and the input disturbances are a disturbance at the plant output with frequency weighting $\hat{W}_1(\lambda)$ and measurement noise with frequency weighting $\hat{W}_2(\lambda)$. That is,

$$\Phi = \begin{pmatrix} (1-PK)^{-1}W_1 & PK(1-PK)^{-1}W_2 \\ \rho K(1-PK)^{-1}W_1 & \rho K(1-PK)^{-1}W_2 \end{pmatrix}$$

where

$$\hat{W}_1(\lambda) = \frac{0.4}{1-0.6\lambda}; \quad \hat{W}_2(\lambda) = \frac{1-0.75\lambda}{1-0.25\lambda}$$

$\rho = 0.1$ and $\hat{P}(\lambda)$ is the same as in Example I. Then, the results in Table V are obtained by applying the above algorithm starting with $N = 3$. For $N = 10$, the suboptimal controller is of order five and achieves a norm that is within half a percent of the optimal. (The jump in order is most likely due to convergence to another optimal solution.) In contrast, it can be shown that the FMV method has no polynomial feasible solution for any N (due to the way $\hat{W}_1$ and $\hat{W}_2$ enter the problem). This example shows how the delay augmentation algorithm can generate low order suboptimal controllers even when the FMV method has no solution.

TABLE V
EXAMPLE II: DELAY AUGMENTATION ALGORITHM

			DA			FMV	
N	$\underline{\mu}_N$	$\overline{\mu}_N$	$\operatorname{len}(\Phi_N^o)^T$	ord(K)	Comments	$\overline{\nu}_N$	ord(K)
3	60.453	102.34	$\begin{pmatrix} 5 & 3 \\ 5 & 3 \end{pmatrix}$	4	Reorder inputs	—	—
3	60.400	81.161	$\begin{pmatrix} 3 & 5 \\ 3 & 5 \end{pmatrix}$	2	Keep order	—	—
4	64.702	81.161	$\begin{pmatrix} 3 & 6 \\ 3 & 7 \end{pmatrix}$	2	"	—	—
5	68.284	81.161	$\begin{pmatrix} 3 & 7 \\ 5 & 9 \end{pmatrix}$	2	"	—	—
6	70.721	72.850	$\begin{pmatrix} 6 & 7 \\ 6 & 11 \end{pmatrix}$	5	"	—	—
7	70.754	71.874	$\begin{pmatrix} 6 & 8 \\ 8 & 13 \end{pmatrix}$	5	"	—	—
8	70.888	71.500	$\begin{pmatrix} 6 & 9 \\ 10 & 15 \end{pmatrix}$	5	"	—	—
9	71.040	71.615	$\begin{pmatrix} 6 & 11 \\ 10 & 17 \end{pmatrix}$	5	"	—	—
10	71.089	71.408	$\begin{pmatrix} 6 & 12 \\ 12 & 19 \end{pmatrix}$	5	"	—	—
11	71.110	71.146	$\begin{pmatrix} 12 & 13 \\ 13 & 21 \end{pmatrix}$	12	"	—	—
12	71.113	71.122	$\begin{pmatrix} 13 & 14 \\ 19 & 23 \end{pmatrix}$	14	"	—	—

IX. SUPPORT STRUCTURE OF OPTIMAL SOLUTIONS

Here we explore the support characteristics of the optimal solution in more general terms. The numerical examples in the previous section suggest that it may be possible to infer the support of the optimal solution by observing how the superoptimal solutions, Φ_N^o, evolve as N increases. Here we make an important step in this direction by showing that such support structure is "hinted to" by the support of the sequence of superoptimal solutions.

We have already shown that, given a multiblock problem, there exists a subsequence of super-optimal dual solutions, $G_{N_i}^o$, whose *weak** limit point, G^o, is feasible and optimal (Theorem 7.5). By exploiting this result in combination with the alignment conditions, we will show that the finitely supported partition of the optimal solution is eventually "captured" by the sequence of superoptimal solutions. For that purpose we need the following well known lemma.

Lemma 9.1: If a sequence $G_N \in l_\infty^{n \times m}$ converges *weak** to G, then for any positive integer $L < \infty$, $\|P_L(G_N - G)\|_\infty \to 0$ as $N \to \infty$.

Note that the above lemma implies that each individual entry of G_N also enjoys this convergence property, i.e., $\|P_L(g_{ij,N} - g_{ij})\|_\infty \to 0$ as $N \to \infty$, for all $i = 1,\cdots,n$ and $j = 1,\cdots,m$.

Next, let us review the alignment properties of the optimal solutions. Optimality implies that each optimal solution to the primal problem must be aligned with every optimal solution to the dual problem. In particular, if an optimal dual solution, G^o, is such that

$$|g_{ij}^o(t)| < \max_{1 \le j \le n_w} \|g_{ij}^o\|_\infty \quad \text{for all } t > T$$

then all optimal primal solutions are such that $\phi_{ij}^o(t) = 0$ for $t > T$. Note that, according to the notation developed in Section V, $\max_{1 \le j \le n_w} \|g_{ij}^o\|_\infty$ is nothing but $\gamma_0^o(i)$. The next theorem puts all these pieces together.

Theorem 9.1: Given a multiblock problem, if all optimal dual solutions are such that $|g_{ij}^o(T)| = \gamma_0^o(i)$ for some $T \in \mathbf{Z}_+$ and $|g_{ij}^o(t)| < \gamma_0^o(i)$ for all $t > T$ then, for every $L > T$ there exists a positive integer N^* such that $\phi_{ij,N}^o(t) = 0$ for $T < t \le L$ and for any $N \ge N^*$.

Proof: (Note: to simplify notation we drop subindexes i, j and superindex 'o'.) Given some $L > T$, pick $\epsilon > 0$ such that

$$\min_{T < t \le L} (\gamma_0 - |g(t)|) = \epsilon. \tag{44}$$

By Lemma 9.1, for every $L > T$ there exists N^* such that

$$\|P_L(g_N - g)\|_\infty < \frac{\epsilon}{2} \tag{45}$$

for all $N > N^*$. First we prove (by contradiction) that $|g_N(t)| < \gamma_{o,N}$ for $T < t < L$ and for any $N \geq N^*$. The result then follows from the alignment conditions.

Given $N > N^*$, assume that $|g_N(t_1)| = \gamma_{o,N}$ for $T < t_1 \leq L$. Then, by (44) and (45),

$$\gamma_{0,N} - \gamma_0 \leq |g_N(t_1)| - |g(t_1)| - \epsilon < \frac{\epsilon}{2} - \epsilon.$$

Therefore,

$$\gamma_0 - \gamma_{0,N} > \frac{\epsilon}{2}. \tag{46}$$

Next, consider the point $t = T$. From (45) and the fact that $|g_N(t)| \leq \gamma_{0,N}$ in general, we have

$$\gamma_0 - \gamma_{0,N} \leq |g(T)| - |g_N(T)| < \frac{\epsilon}{2}$$

which contradicts (46). This implies that $\phi_N(t) = 0$ for $T < t \leq L$ and $N > N^*$ which is the desired result. ∎

In other words, given the conditions of the theorem above, and for N large enough, there is a "gap" of zeros (between T and L) in $\phi^o_{ij,N}(t)$ which gets wider as N increases, i.e., as L increases. However, T does not change for N large enough, giving a clue on the length of the finitely supported entries of Φ^o. The difficulty is that we do not have an *a priori* estimate of how large N has to be to capture T.

It is worth pointing out that Theorem 9.1 can be applied to the FMV sequence of suboptimal solutions too, since the corresponding duals also have a *weak** convergent subsequence [34]. However, there is an important difference in the way the DA and FMV sequence of solutions behave, which was pointed out in the previous section. Indeed, while the FMV solutions are consistently supported for $t > L$, the DA solutions are not. This observation was crucial in constructing low order suboptimal controller. We expand these ideas in the following section.

X. Observations

This section includes a few observations based on a fair amount of computational experience using the delay augmentation method and on some intuitive ideas on the problem of l_1 optimization in general. It is by no means a formal or precise presentation. It is simply intended to give some lead into new ideas that might open the way to finding the exact solution of multiblock problems in general. In particular, a conjecture is stated, establishing a stronger connection between the support structure of the optimal solution and the DA method.

Observe the way the DA method works. It transforms a general multiblock problem into a square one, therefore generating polynomial superoptimal solutions, Φ^o_N. Without changing the order of inputs and outputs, the sequence Φ^o_N will increase its length as N increases. However, it was noted in previous examples that not every entry of Φ^o_N increases its length in the same way. In fact, a closer look at the sequence Φ^o_N suggests that the support of some of its entries stops changing after some N. This is exactly what happened in Example I, cases a) and b), where the support of one of the entries of Φ^o_N remained the same after some N regardless of the ordering. In Example II, the pattern also occurs but for $N > 12$ (not shown in Table V). Next, note that $\Phi^o_{11,N} = \overline{\Phi}_{11,N}$ since that block of the problem is not affected by the extra free parameters. Therefore, for each N, $\overline{\Phi}_{11,N}$ is polynomial. Then, if those entries of Φ^o_N that have constant support after some N are collected (by reordering) in $\Phi^o_{11,N}$, $\overline{\Phi}_{11,N}$ will have constant support. Interestingly, those entries of constant support seem to be always enough to define a one-block partition and therefore fill the necessary entries of $\overline{\Phi}_{11,N}$. Furthermore, many multiblock problems seem to have this property.

A multiblock problem in this class can be viewed as dominated by a one-block partition. In other words, there is an embedded one-block problem that is further constrained by the rank interpolation conditions. Such constraints, however, are not enough to change the polynomial nature of the optimal solution corresponding to that partition, although, in general, they have the effect of increasing its order. With this we extend the notion of TD one-block partitions where the added constraints due to the rank interpolation conditions were totally inactive.

Definition 10.1: Given a multiblock problem, a one-block partition is partially dominant (PD) if all l_1 optimal solutions are polynomial in the entries corresponding to such partition.

Clearly, a TD one-block partition is also PD but not vice versa. Based on this definition we state the following conjecture.

Conjecture 10.1: Given a multiblock problem with a PD one-block partition, there exists a positive integer N^* such that the DA solution, Φ^o_N, for $N > N^*$ captures the exact support of the sequences corresponding to the PD one-block partition. Furthermore, since the actual linear program splits Φ^o_N into the difference of two positive sequences (Φ^{o+}_N and Φ^{o-}_N), the sign of the nonzero entries of the exact solution corresponding to the PD partition is also captured. That is, for any pair of indexes i, j) in the PD partition, and $N > N^*$,

$$\phi^o_{ij}(k) = 0 \Leftrightarrow \phi^o_{ij,N}(k) = 0$$

$$\phi^o_{ij}(k) > 0 \Leftrightarrow \phi^o_{ij,N}(k) > 0$$

$$\phi^o_{ij}(k) < 0 \Leftrightarrow \phi^o_{ij,N}(k) < 0.$$

This conjecture is supported by a fair amount of numerical experiments covering the most obvious combinations (i.e., two-block row and column problems and four-block problems with different input–output dimensions). At the same time, it is consistent with Theorem 9.1 but stronger. Indeed, the conjecture claims that the superoptimal solution will not be supported for $t \geq L$. This conjecture, if proven correct, has interesting consequences. To illustrate some of the ideas involved, consider the following simple

two-block column problem:

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} - \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} q$$

and assume, without loss of generality, that $\hat{u}_1$ and $\hat{u}_2$ are polynomials (this can always be obtained by polynomial factorization of $\hat{U}$). Further, assume that $(h_1 \quad h_2)^T$ is a polynomial feasible solution and that the outputs are ordered such that ϕ_1 is PD. Then we have the following equality due to the rank interpolation conditions:

$$\hat{u}_2 \hat{\phi}_1^o - \hat{u}_1 \hat{\phi}_2^o = \hat{u}_2 \hat{h}_1 - \hat{u}_1 \hat{h}_2. \tag{47}$$

Assume that all zeros common to $\hat{u}_1$ and $\hat{u}_2$ have been canceled out from the above equation. Clearly, the right-hand side of (47) is polynomial, and furthermore, the first term on the left-hand side is polynomial since we assumed that ϕ_1 is PD. Therefore, the second term on the left-hand side must be polynomial. This implies that two situations are possible: either $\hat{\phi}_2^o$ is polynomial or it has stable poles that are canceled by stable zeros of $\hat{u}_1$.

This observation has interesting implications. On one hand, there is a class of multiblock problems with polynomial optimal solutions that is characterized by the absence of stable zeros in $\hat{u}_1$. Such solutions can then be computed exactly by either the FMV or the DA method. On the other hand, if $\hat{u}_1$ has stable zeros and ϕ_2^o is infinitely supported, the rate at which ϕ_2^o decays is given by a subset of the stable zeros of $\hat{u}_1$. This information could be used to transform the original problem into a finite dimensional one for which exact solutions are computable. This approach is currently under investigation. It should be noted that the above ideas can be easily extended to the general multiblock problem.

Finally, note that if the above conjecture is correct, the DA algorithm would automatically reorder any TD partition in Φ_{11} and provide the exact answer, without the need to solve all possible combinations of one-block problems (see discussion after Corollary 7.3).

XI. A Synthesis Example

In this section, we apply the DA method to a specific control problem, namely, the pitch axis control of the X29 aircraft. The motivation for doing so is two-fold: first, to illustrate the use of the delay augmentation method in a more realistic problem, and second, to have a first look at the frequency domain features of an l_1-optimal design (albeit for one particular example). In order to give some perspective to this presentation, we will compare the characteristics of the l_1 design with those of an $\mathscr{H}_\infty$ optimal design.

It should be stressed, however, that this particular control problem was not chosen for the purpose of demonstrating extreme behaviors of the l_1 and $\mathscr{H}_\infty$ optimal solutions. Rather, it was candidly selected as an interesting control problem in general.

The X29 aircraft poses an interesting control problem due to its revolutionary forward-swept wing design. With such configuration, the center of gravity lies behind the aerodynamic center of pressure, rendering the aircraft statically unstable. Thus, a control system has to actively stabilize the aircraft during flight.

We are interested in designing a digital controller for a simple model of the pitch dynamics of the aircraft. The airplane has three types of control surfaces: canard wings, flaperons on the main wings and strakes on the tail. In order to simplify the model, the action of these control surfaces are lumped into one equivalent actuator with first order dynamics. Similarly, the gyroscopes and accelerometers are modeled by an equivalent sensor with neglectable dynamics. Thus, the system can be approximately represented by the following continuous time SISO plant [35]:

$$\hat{P}(s) = \underbrace{\frac{(s+3)}{(s+10)(s-6)}}_{\text{airframe}} \quad \underbrace{\frac{20}{(s+20)}}_{\text{equiv. actuator}} \quad \underbrace{\frac{(s-26)}{(s+26)}}_{\text{overhead}} \tag{48}$$

where s is the Laplace variable. The airframe factor corresponds to a simplified model of the pitch dynamics of the airplane flying at a low altitude and with an air speed of approximately 0.9 Mach. The overhead factor lumps the equivalent low frequency phase lag introduced by the dynamics that were neglected in deriving the reduced model (48). In particular, this all-pass factor is an approximate representation of the collected phase lag of the gyroscopic sensor dynamics, the actuator servo dynamics, the airframe flexible modes, and the digital implementation (i.e., pre-filter, zero order hold and computing delay) corresponding to a sampling period $\Delta t = 1/30$ seconds.

Consider the following formal synthesis problem:

$$\inf_{K \text{ stab}} \left\| \begin{matrix} W_1 KS \\ W_2 S \end{matrix} \right\|_1$$

where S is the sensitivity function. Such problem requires the discrete time version of (48) and two weighting transfer functions. The λ-domain model of the plant, $\hat{P}(\lambda)$, is obtained by discretizing (48) assuming a zero order hold at the plant input and a synchronized sampling of the (pre-filtered) plant output. The weights are chosen as follows: let $\hat{W}_1$ be a scalar equal to 0.01 and let $\hat{W}_2(\lambda)$ be the discrete time version of the continuous time transfer function $(s+1)/(s+0.001)$ for a sampling period $\Delta t = 1/30$. This choice of weights reflects a trade-off between low frequency performance and the control effort.

Note that a controller designed for the discrete-time model of a continuous-time plant completely ignores the inter-sampling behavior of the system. An optimal controller designed in this way is actually suboptimal for the original hybrid system. This notwithstanding, we will carry out the design and comparison entirely in the discrete

domain (both for l_1 and $\mathscr{H}_\infty$ designs), taking the discrete time plant model and weights as the starting point.

A. Computing an l_1-Suboptimal Controller

With this problem set-up we are ready to apply the delay augmentation algorithm as described in Section IX. Table VI shows the sequence of results obtained in this case, starting with $N = 4$. Note how the length of the response corresponding to the weighted sensitivity stops increasing after $N = 7$, suggesting that such row is PD. For $N = 80$ the achieved l_1 norm is within one percent of the optimal so we stop the iteration process. It is interesting to note how slowly the upper bound converges to the optimal. This behavior is consistent with the observations made in Section X regarding the rate of decay of Φ when one row is PD. Indeed, if the first row corresponding to the weighted sensitivity is PD, then the rate of decay of the second row is dictated by the stable zeros of $\hat{u}_1(\lambda)$. It is easy to check that such transfer function contains two stable zeros that are close to the unit circle. Then, if the optimal second row decays slowly, the extra free parameter (q_2) corresponding to the DA solution will be significant even for large values of N.

Next, we will compare the time and frequency domain characteristics of the l_1 suboptimal design corresponding to $N = 80$ with an $\mathscr{H}_\infty$ design. The comparison will be based on three different aspects of the solutions: 1) operator norms, 2) frequency response characteristics, and 3) time response characteristics.

Table VII shows how the l_1 and $\mathscr{H}_\infty$ norms of the two solutions compare. As expected, the $\mathscr{H}_\infty$ design achieves better $\mathscr{H}_\infty$ norms while the l_1 design achieves better l_1 norms. A cross examination shows that both solutions are fairly good in terms of both measures. In fact, this does not come as a surprise in view of the following norm inequality [5] valid for any stable finite dimensional system $H \in l_1^{p \times q}$:

$$\|\hat{H}\|_{\mathscr{H}_\infty} \leq \sqrt{p}\|H\|_1 \leq \sqrt{p}(2n+1)\sqrt{q}\|\hat{H}\|_{\mathscr{H}_\infty}$$

where n is its McMillan degree. Thus, minimizing any of the two norms will also "push down" the other one, particularly in a low order problem as the one under consideration.

Next, let us examine the frequency domain features. Both designs have failry similar frequency domain characteristics as shown in Figs. 2 and 3. While the l_1 design has better disturbance rejection at low and medium frequencies, it overshoots at high frequencies where the $\mathscr{H}_\infty$ norm is achieved. In fact, Fig. 3 shows that both controllers have very similar response, the only significant difference being at frequencies close to $\pi/\Delta t$. An interesting difference, though, is that the l_1 design results in an unstable controller while the $\mathscr{H}_\infty$ design does not. Finally, we compare the weighted and unweighted sensitivity step response of both designs (Figs. 4 and 5). Note how the output of the plant, y, converges to zero faster in the l_1 design than in the $\mathscr{H}_\infty$ design (Fig. 5). This is a direct result of the smaller weighted steady state error in the l_1 design (see Fig. 4) and the pole of $\hat{W}_2$ at 0.9999 (almost a pure integrator).

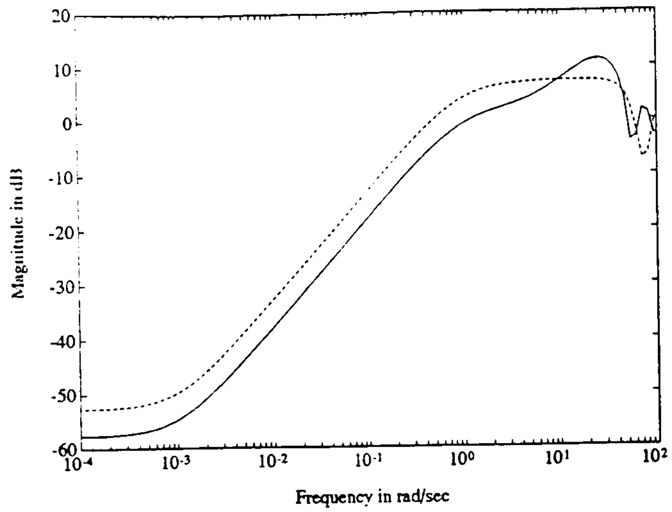


Fig. 2. Frequency response of $\hat{S}$ for l_1 design (full line) and $\mathscr{H}_\infty$ design (dashed line).

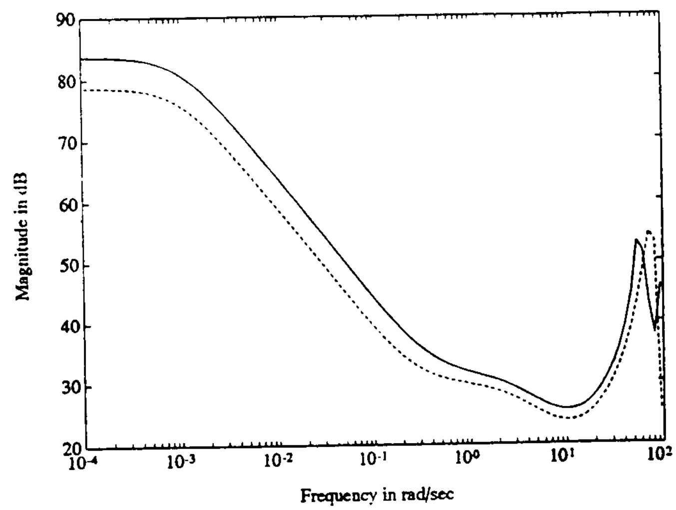


Fig. 3. Frequency response of $\hat{K}$ for l_1 design (full line) and $\mathscr{H}_\infty$ design (dashed line).

TABLE VI
X29 SYNTHESIS PROBLEM: DELAY AUGMENTATION ALGORITHM

N	$\underline{\mu}_N$	$\overline{\mu}_N$	len $(\Phi^o_N)^T$	ord(K)	Comments
4	3.254	1256.4	(10 5)	11	Reorder outputs
5	4.024	7.619	(5 5)	6	Keep order
6	4.045	5.059	(5 6)	6	"
7	4.048	5.052	(6 7)	6	"
8	4.051	4.652	(6 8)	6	"
9	4.051	4.319	(6 9)	6	"
10	4.052	4.224	(6 10)	6	"
⋮					
20	4.053	4.196	(6 20)	6	"
⋮					
40	4.053	4.158	(6 36)	6	"
⋮					
80	4.054	4.091	(6 69)	6	"

XII. CONCLUSIONS

A complete and comprehensive study of the general l_1-optimal multiblock problem has been presented. It ad-

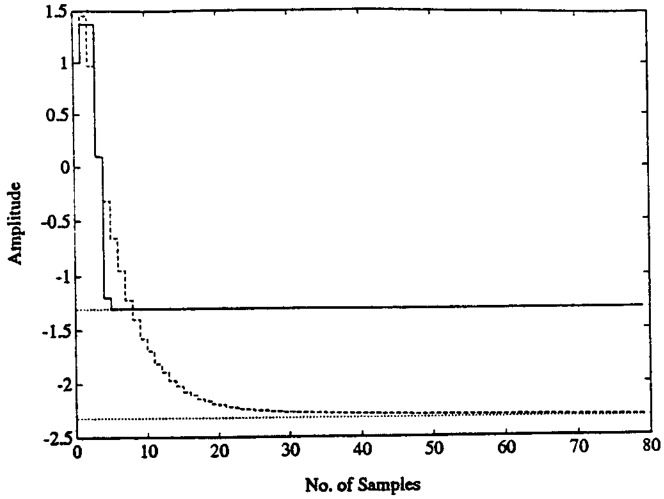


Fig. 4. Weighted sensitivity step response for l_1 design (full line) and $\mathcal{H}_\infty$ design (dashed line).

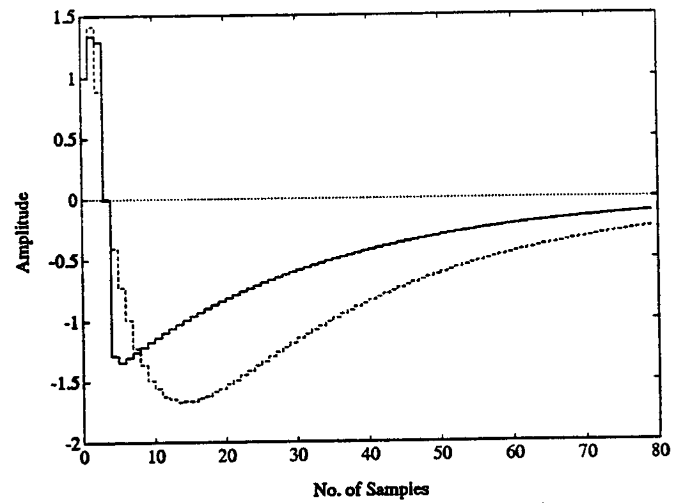


Fig. 5. Sensitivity step response from w to y for l_1 design (full line) and $\mathcal{H}_\infty$ design (dashed line).

TABLE VII
OPERATOR NORM COMPARISON ($\Delta t = 1/30$)

		$\|\cdot\|_{\mathcal{H}_\infty}$	$\|\cdot\|_1$	ord(K)
$\mathcal{H}_\infty$ design	Φ	2.4	5.2	5
	W_1KS	2.0	3.3	
	W_2S	2.4	5.2	
l_1 design	Φ	3.8	4.1	6
	W_1KS	2.8	4.1	
	W_2S	3.6	4.1	

vances the understanding of these problems both from a theoretical and a practical point of view.

The paper makes the following contributions:

1) The interpolation conditions are stated in a concise and natural way. As a result the general theory is developed in simpler terms and with a minimum number of assumptions.

2) Methods for computing the interpolation conditions were tied directly to matrix theory.

3) Further insight was gained on the structure of the optimal solution which allowed us to distinguish between different classes of multiblock problems (i.e., problems with TD or PD one-block partitions).

4) A new method for computing suboptimal (or optimal in some special cases) solutions was proposed that exploits such structure. With this method, a sequence of suboptimal controllers can be computed iteratively avoiding (for a class of problems) the problem of order inflation. Each iteration requires the solution of one finite dimensional linear program, and generates upper and lower bounds of the optimal norm with the proper convergence properties. In contrast, previously known approximation schemes required the solution of two linear programs at each iteration, and generated suboptimal controllers with increasing order. In addition, the DA method unifies the treatment of zero and rank interpolations and avoids the coprime factorization of U and V (this was required in previous work [29]). Further, this approach generates a minimal set of constraints describing the feasible subspace [18].

5) A result was presented relating the support characteristics of the optimal and superoptimal solutions, followed by a stronger conjecture.

Several examples were worked out to illustrate the properties of the DA method. In particular, a multiblock problem corresponding to the X29 pitch axis control was solved. The operator norms and frequency domain properties of the solutions were compared with those of a standard $\mathcal{H}_\infty$ design. Although the designs turned out to be quite similar, some differences were found at high frequencies.

As a final note, let us point out that there are still important open questions to be answered in connection with l_1 optimization. From a theoretical point of view, stronger results regarding the support structure of the optimal solution are needed. In particular, a proof or a counter example for the conjecture presented. As pointed out before, proving such conjecture could provide the insight to uncover the underlying finite dimensional structure that the general multiblock problem may have. Also, the existence in general of optimal rational solutions is an interesting open question connected to the above.

Finally, a model reduction theory in the context of l_1 optimization would be of significant practical value. Recall that multiblock as well as one-block problems may have high order optimal controllers (depending on the interpolation data). A straightforward approach to computing lower order suboptimal controllers results from restricting the appropriate entries of Φ to have fixed finite support. But such approach may be far from optimal. Therefore, optimal model reduction techniques would be useful in practical design.

REFERENCES

[1] J. A. Ball, I. Gohberg, and L. Rodman, "Realization and interpolation of rational matrix functions," *Operator Theory: Advances Appl.*, vol. 33, 1988.

[2] J. A. Ball and M. Rakowski, "Zero-pole structure of nonregular

rational matrix functions," *Operator Theory: Advances Appl.*, vol. 47, 1990.

[3] B. A. Bamieh, M. A. Dahleh, and J. B. Pearson, "Minimization of the L^{∞}-induced norm for sampled-data systems," *IEEE Trans. Automat. Contr.*, vol. 38, no. 5, May 1993.

[4] J. S. Baras, D. C. MacEnany, and R. L. Munach, "Fast error-free algorithms for polynomial matrix computations," in *Proc. 29th Conf. Decision Contr.*, Hawaii, 1990.

[5] S. P. Boyd and J. C. Doyle, "Comparison of peak and RMS gains for discrete-time systems," *Syst. Contr. Lett.*, vol. 9, pp. 1–6, 1987.

[6] M. A. Dahleh, "BIBO stability robustness in the presence of coprime factor perturbations," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 352–355, 1992.

[7] M. Dahleh and M. A. Dahleh, "Optimal rejection of persistent and bounded disturbances: continuity properties and adaptation," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 687–696, 1990.

[8] M. A. Dahleh and M. H. Khammash, "Controller design for plants with structured uncertainty," *Automatica*, vol. 29, no. 1, Jan. 1993.

[9] M. A. Dahleh and J. B. Pearson, "l_1-Optimal feedback controllers for MIMO discrete-time systems," *IEEE Trans. Automat. Contr.*, vol. 32, pp. 314–322, 1987.

[10] ——, "L^1 Optimal compensators for continuous-time systems," *IEEE Trans. Automat. Contr.*, vol. 32, pp. 889–895, 1987.

[11] ——, "Optimal rejection of persistent disturbances, robust stability and mixed sensitivity minimization," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 722–731, 1988.

[12] ——, "Minimization of a regulated response to a fixed input," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 924–930, 1988.

[13] M. A. Dahleh and Y. Ohta, "A necessary and sufficient condition for robust BIBO stability," *Syst. Contr. Lett.*, vol. 11, pp. 271–275, 1988.

[14] M. A. Dahleh, P. Voulgaris, and L. Valavani, "Optimal and robust controllers for periodic and multirate systems," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 90–99, 1992.

[15] G. Deodhare and M. Vidyasagar, "Some results on l_1-optimality of feedback control systems: The SISO discrete-time case," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 1082–1085, 1990.

[16] C. A. Desoer and J. D. Schulman, "Zeros and poles of matrix transfer functions and their dynamical interpretation," *IEEE Trans. C-S*, vol. 21, pp. 3–8, 1974.

[17] I. J. Diaz-Bobillo and M. A. Dahleh, "State feedback l_1 optimal controllers can be dynamic," *Syst. Contr. Lett.*, vol. 19, pp. 87–93, 1992.

[18] I. J. Diaz-Bobillo, "The general l_1 optimal multiblock problem: Exact and approximate solutions," Ph.D. dissertation, Tech. Rep. LIDS-TH-2093, Massachusetts Institute of Technology, Feb. 1992.

[19] G. Dullerud and B. A. Franics, "$\mathcal{L}^1$ Analysis and design of sampled-data systems," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 436–446, 1992.

[20] A. Emami-Naeini and P. Van Dooren, "Computation of zeros of linear multivariable systems," *Automatica*, vol. 14–4, pp. 415–430, 1982.

[21] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*. New York: Academic, 1982.

[22] G. H. Golub and C. F. Van Loan, *Matrix Computations*. New York: John Hopkins Univ. Press, 1983.

[23] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[24] M. Khammash and J. B. Pearson, "Performance robustness of discrete-time systems with structured uncertainty," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 398–412, 1991.

[25] E. Kreyszig, *Introductory Functional Analysis with Applications*. New York: Wiley, 1989.

[26] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.

[27] ——, *Linear and Nonlinear Programming*. New York: Addison-Wesley, 1984.

[28] A. G. MacFarlane and N. Karcanias, "Poles and zeros of linear multivariable systems: A survey of the algebraic, geometric and complex-variable theory," *Int. J. Contr.*, vol. 24, pp. 33–74, 1976.

[29] J. S. McDonald and J. B. Pearson, "l_1-Optimal control of multivariable systems with output norm constraints," *Automatica*, vol. 27, pp. 317–329, 1991.

[30] M. A. Mendlovitz, "A simple solution to the l_1 optimization problem," *Syst. Contr. Lett.*, vol. 12, pp. 461–463, 1989.

[31] C. B. Schrader and M. K. Sain, "Research on system zeros: A survey," *Int. J. Contr.*, vol. 50, pp. 1407–1433, 1989.

[32] J. S. Shamma and M. A. Dahleh, "Time-varying versus Time-invariant compensation for rejection of persistent bounded disturbances and robust stabilization," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 838–847, 1991.

[33] O. J. Staffans, "Mixed sensitivity minimization problems with rational l_1-optimal solutions," *J. Optimiz. Theory Appl.*, vol. 70, pp. 173–189, 1991.

[34] ——, "On the four-block model matching problem in l_1," Helsinki Univ. Technology, Espoo, Tech. Rep. A289, 1990.

[35] G. Stein, $\mathcal{H}_\infty$ Workshop, Session II, in *Proc. IEEE Conf. Decision Contr.*, 1987.

[36] P. M. Van Dooren, P. Dewilde, and J. Vandewalle, "On the determination of the Smith–McMillan form of a rational matrix from its Laurent expansion," *IEEE Trans. C-S*, vol. 26, 1979.

[37] M. Vidyasagar, "Optimal rejection of persistent bounded disturbances," *IEEE Trans. Automat. Contr.*, vol. 31, pp. 527–534, 1986.

[38] D. C. Youla, H. A. Jabr, and J. J. Bongiorno, "Modern Wiener–Hopf design of optimal controllers-Part 2: The multivariable case," *IEEE Trans. Automat. Contr.*, vol. 21, pp. 319–338, 1976.

Ignacio J. Diaz-Bobillo (M'92) was born in Buenos Aires, Argentina. He received the electrical engineering degree from the Buenos Aires Institute of Technology, Buenos Aires, Argentina, in 1981, and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, MA, in 1985 and 1992, respectively.

From 1981 to 1988, he held a research and teaching position at the National Commission of Space Research in Buenos Aires, Argentina. Currently, he is a Postdoctoral Associate at the Laboratory for Information and Decision Systems in the Department of Electrical Engineering and Computer Science of the Massachusetts Institute of Technology. His general areas of interest include systems theory, robust control and identification.

Dr. Diaz-Bobillo is the co-author (with Munther A. Dahleh) of the book *Control of Uncertain Systems: A Linear Programming Approach*, to be published by Prentice-Hall.

Munther A. Dahleh (S'84–M'87) was born in 1962. He received the B.S. degree from Texas A & M University, College Station, TX, in 1983, and the Ph.D. degree from Rice University, Houston, TX, in 1987, all in electrical engineering.

Since 1987, he has been with the Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, where he is now a Finmeccanica Associate Professor. He has been a visiting Professor at the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA, for the Spring of 1993. He has held consulting positions with C.S. Draper Laboratory and NASA since 1988, and FIAT Research Center since 1991. His interests include robust control and identification, application of mathematical programming in synthesizing multivariable controllers, and the analysis and design of Hybrid systems.

Dr. Dahleh has been the recipient of the Raplh Budd award in 1987 for the best thesis at Rice University, George Axelby outstanding paper award (paper coauthored with J. B. Pearson in 1987), an NSF presidential young investigator award (1991), the Finmeccanica career development chair (1992) and the Donald P. Eckman Award of the American Control Council. He is currently serving as an Associate Editor for the Systems and Control Letters and IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is the author (with Ignacio J. Diaz-Bobillo) of the book *Control of Uncertain Systems: A Linear Programming Approach*, to be published by Prentice-Hall.

Systems & Control Letters 20 (1993) 157–166
North-Holland

The sample complexity of worst-case identification of FIR linear systems *

Munther A. Dahleh, Theodore V. Theodosopoulos and John N. Tsitsiklis

Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 1 May 1992
Revised 7 October 1992

Abstract: We consider the problem of identification of linear systems in the presence of measurement noise which is unknown but bounded in magnitude by some $\delta > 0$. We focus on the case of linear systems with a finite impulse response. It is known that the optimal identification error is related (within a factor of 2) to the diameter of a so-called uncertainty set and that the latter diameter is upper-bounded by 2δ, if a sufficiently long identification experiment is performed. We establish that, for any $K \geq 1$, the minimal length of an identification experiment that is guaranteed to lead to a diameter bounded by $2K\delta$ behaves like $2^{Nf(1/K)}$, when N is large, where N is the length of the impulse response and f is a positive function known in closed form. While the framework is entirely deterministic, our results are proved using probabilistic tools.

Keywords: Worst-case identification; sample complexity; bounded but unknown disturbance.

1. Introduction

Recently, there has been increasing interest in the problem of worst-case identification in the presence of bounded noise. In such a formulation, a plant is known to belong to a model set $\mathscr{M}$, and its measured output is subject to an unknown but bounded disturbance. The objective is to use input/output information to derive a plant estimate that approximates the true plant as closely as possible, in some induced norm. For frequency domain experiments, algorithms that guarantee accurate identification in the $\mathscr{H}_\infty$ setting were furnished in [4,5,6,7]. For general experiments, algorithms that guarantee accurate identification in the ℓ_1 sense were suggested in [17,18]. These algorithms are based on the Occam's Razor principle by which the simplest model is always used to explain the given data. The optimal asymptotic worst-case error is characterized in terms of the diameter of the 'uncertainty set': the set of all plants consistent with all the data and the noise model. Other related work on the worst-case identification problem can be found in [8,10,11,19]. In particular, [10] presents a specific experiment that uses a Galois sequence as an input, and shows that the standard Chebyshev algorithm results in an asymptotic error bounded by the worst-case diameter of the uncertainty set. A Galois sequence is constructed by concatenating a countable number of finite sequences, such that the k-th sequence contains all possible combinations of $\{-1, +1\}$ of length k, and so it is rich enough to accurately identify exactly k parameters of the impulse response. The length of each sequence is clearly exponential in k. Finally, identification problems with bounded but unknown noise were studied in the context of prediction (not worst-case) in [12,13]. Other related work, for nonlinear systems, can be found in [3].

An important result from the work of [17,18] states that for the model set of all stable plants, accurate identification in the ℓ_1 sense is possible if and only if the input excites all possible frequencies on the unit circle. This is due to two reasons: the first is that bounded noise is quite rich and the second is that

Correspondence to: Prof. M.A. Dahleh, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

* Research supported by the AFOSR under grant AFOSR-91-0368, and by the NSF under grants 9157306-ECS and ECS-8552419, and by a grant from Siemens AG.

minimizing an induced norm such as the ℓ_1 norm implies that the estimate has a very good predictive power. Inputs with such properties tend to be quite long, and this suggests that the sample complexity of this kind of identification problems tends to be quite high, as a function of the numbers of estimated parameters of the impulse response.

In this paper, we will study the sample complexity (required length) of the inputs for worst-case identification of FIR plants, under the ℓ_1 norm, in the presence of arbitrary bounded measurement noise. It will be shown that in order to guarantee that the diameter of the uncertainty set is bounded by $2K\delta$, where δ is the bound on the noise and K is a constant (larger than 1), the length of the input must increase like $2^{Nf(1/K)}$, where N is the length of the impulse response and f is a positive function. Since the worst-case error is at least half of the diameter, these results show that the sample complexity is exponential in N even if the allowable accuracy is far from optimal, and capture the limitations of accurate identification in the worst-case set-up. We also show that our sample complexity estimate is tight, in the sense that there exist inputs of length approximately equal to $2^{Nf(1/K)}$ that lead to a $2K\delta$ bound on the diameter. An interesting technical aspect of this paper is that the existence of such inputs is established by means of a probabilistic argument reminiscent of the methods commonly employed in information theory.

Other researchers have also recently addressed the sample complexity of worst-case identification. In a personal discussion with Poolla (January 1992), he pointed out to us (specifically to Dahleh) that the optimal identification case had exponential complexity, as in the lower bound of our Theorem 2.1. We have recently received a preprint by Poolla and Tikku [14] which, among other results, contains exponential lower bounds for the sample complexity of suboptimal identification of FIR systems. These lower bounds are similar to, although somewhat weaker than, the lower bound in our Theorem 2.2. Chronologically, the results of [14] precede ours, although we didn't have knowledge of their results when writing our paper. Finally, [14] contains some upper bounds but, unlike our Theorem 2.2, they are far from being tight. Also, while writing our paper, we learned that Milanese [9] had arrived to results similar to the exponential lower bound in our Theorem 2.1. His report does not contain any discussion of the case where the error is within a factor of the optimal.

2. Problem definition

Let $\mathcal{M}_N$ be the set of all linear systems with a finite impulse response of length N. Any element h of $\mathcal{M}_N$ will be identified with a finite sequence $(h_1,\dots,h_N)\in\mathbb{R}^N$. Let U_n be the set of all infinite real sequences $\{u_i\}_{i=1}^{\infty}$ such that $|u_i|\le 1$ for all i, and $u_i=0$ for $i>n$. Any element of U_n will be called an *input of length n*. Finally, for any positive number δ, let D_δ, called the *disturbance set*, be the set of all infinite sequences $d=\{d_i\}_{i=1}^{\infty}$ such that $|d_i|\le\delta$ for all i.

We are interested in experiments of the following type: an input $u\in U_n$ is applied to an (unknown) system $h\in\mathcal{M}_N$, and we observe the noisy measurement

$$y=h*u+d, \tag{2.1}$$

where $*$ denotes convolution, and where $d\in D_\delta$ plays the role of an output disturbance or measurement noise. It is clear that, for $i>N+n$, we have $y_i=d_i$, and y_i carries no useful information on the unknown system h.

The set that contains all plants in the model set that are consistent with the input/output data and the noise model is called the uncertainty set and is given by

$$S_{N,n}(y,u)=\{\phi\in\mathcal{M}_N \mid \|y-\phi*u\|_\infty\le\delta\}$$

The diameter diam(S) of a subset S of ℓ_1 is defined by

$$\operatorname{diam}(S)=\sup_{x,y\in S}\|x-y\|_1.$$

We then define the worst case diameter for a given input $u \in U_n$ by

$$D_{N,n}(u) = \sup_{d \in D_\delta} \sup_{\phi \in \mathscr{M}_N} \operatorname{diam}(S_{N,n}(u * \phi + d, u)).$$

Any identification algorithm that lets its plant estimate be an element of the uncertainty set has an error upper-bounded by the diameter of the uncertainty set. Besides, it is shown in [15,16,17] that the error of any identification algorithm is lower-bounded by half the diameter of the uncertainty set. Define

$$D^*_{N,n} = \inf_{u \in U_n} D_{N,n}(u).$$

It is shown in [17] that

$$\lim_{n \to \infty} D^*_{N,n} = 2\delta. \tag{2.2}$$

Thus, as the length of the experiments increases, and with a suitable identification algorithm, the worst-case error can be made as small as twice the disturbance bound δ, but no smaller than δ. A question that immediately arises is how long should n be for the error to approach 2δ. We address this question by focusing on the behavior of the diameter of the uncertainty set, as the inputs are allowed to become longer.

Let us define

$$n^*(N) = \min\{n \mid D^*_{N,n} = 2\delta\}. \tag{2.3}$$

It is far from a priori clear whether $n^*(N)$ is finite. This is answered by the following theorem which also serves as motivation for the main theorem (Theorem 2.2) of this paper.

Theorem 2.1. [1] *For any* $\delta > 0$ *and* N, *we have* $2^{N-1} + N - 1 \le n^*(N) \le 2^N + N - 1$.

Proof. We start by proving the lower bound on $n^*(N)$. Fix N and let us denote $n^*(N)$ by m. Suppose that $m < \infty$, and let $\mathscr{A}$, $u \in U_m$, be such that $D_{N,m}(u) = 2\delta$. Let $v \in \{-1, 1\}^m$ be defined by $v_i = 1$ if $u_i \ge 0$, and $v_i = -1$ if $u_i < 0$. For notational convenience, we define $u_i = 0$ for $i \le 0$. We distinguish two cases:

(a) Suppose that for every $\phi \in \{-1, 1\}^N$, there eixsts some $i(\phi) \in \{1, \dots, m - N + 1\}$ such that either ϕ or $-\phi$ is equal to $(v_{i(\phi)+N-1}, v_{i(\phi)+N-2}, \dots, v_{i(\phi)})$. It is clear that $i(\phi)$ can be the same for at most two different values of ϕ. Since the number of different choices for ϕ is 2^N, it follows that $m - N + 1 \ge 2^{N-1}$, which proves that $m \ge 2^{N-1} + N - 1$.

(b) Suppose now that the assumption of case (a) fails to hold. Let $\phi \in \{-1, 1\}^N$ be such that both ϕ and $-\phi$ are different from $(v_{i+N-1}, v_{i+N-2}, \dots, v_i)$, for all $i \in \{1, \dots, m - N + 1\}$. Suppose that $h = \delta\phi/(N-1)$. Then

$$|(h * u)_i| = \left| \sum_{k=1}^{N} h_k u_{i-k} \right| = \frac{\delta}{N-1} \left| \sum_{k=1}^{N} \phi_k u_{i-k} \right|. \tag{2.4}$$

Since $|\phi_k| = 1$ and $|u_{i-k}| \le 1$, we see that $|\sum_{k=1}^{N} \phi_k u_{i-k}| \le N$. Let i be such that $N < i \le m$. By our assumption on ϕ, the signs of u_{i-k} cannot be the same as the signs of of ϕ_k for all k, neither the same as the signs of $-\phi_k$ for all k, and this leads to the stronger inequality

$$\left| \sum_{k=1}^{N} \phi_k u_{i-k} \right| \le N - 1. \tag{2.5}$$

[1] We acknowledge Professor Poolla for pointing out an error in the previous version of this theorem.

We finally note that for $i \notin (N, m]$, at least one of the summands $\phi_k u_{i-k}$ is equal to zero, which implies that (2.5) is valid for all i. Combining (2.4) and (2.5), we conclude that $|(h * u)_i| \leq \delta$ for all i. Therefore, there exists a choice for the disturbance sequence d under which the observed output $h * u + d$ is equal to zero at all times. Using the same argument, we see that if $h = -\delta\phi/(N-1)$, there also exists another choice of the disturbance sequence for which the observed output is zero at all times.

We have thus shown that it is possible to observe an output sequence which is identically equal to zero while the true system can be either $\delta\phi/(N-1)$ or $-\delta\phi/(N-1)$. This implies that the worst case diameter satisfies

$$D_{N,m}(u) \geq 2 \| \delta\phi/(N-1) \|_1 > 2\delta. \tag{2.6}$$

But this contradicts the definition of $m = n^*(N)$ and shows that case (*b*) is not possible. Thus, case (a) is the only possible one, and the lower bound has already been established for that case. The upper bound follows easily by using the input sequence proposed in [10,17]. Let u be a finite sequence whose entries belong to $\{-1, 1\}$ and such that for every $\phi \in \{-1, 1\}^N$ there exists some $i(\phi)$ such that $\phi = (u_{i(\phi)}, u_{i(\phi)+1}, \dots, u_{i(\phi)+N-1})$. Such a sequence, called a Galois sequence, can be chosen so that its length is equal to $2^N + N - 1$ [10]. With this input, the worst case diameter is equal to 2δ. □

Theorem 2.1 has the disappointing conclusion that the worst-case error is guaranteed to become at most 2δ only if a very long experiment is performed. In practice, values of N of the order of 20 or 30 often arise. For such cases, the required length of an identification experiment is prohibitively long if an error guarantee as small as 2δ is desired. This motivates the problem studied in this paper: if the objective is to obtain an identification error within a factor K of the optimal value, can this be accomplished with substantially smaller experiments? Theorem 2.2 below is equally disappointing with Theorem 2.1: it shows that experiments of length exponential in N are required to obtain such an error guarantee. The exponent depends of course on K and we are able to compute its asymptotic value (as N increases) exactly.

Theorem 2.2. *Fix some* $K > 1$ *and let*

$$n^*(N, K) = \min\{n \mid D^*_{N,n} \leq 2K\delta\}. \tag{2.7}$$

Then:
(a) $n^*(N, K) \geq 2^{Nf(1/K)-1} - N + 2\lceil N/K \rceil - 1$.
(b) $\lim_{N\to\infty}(1/N)\log n^*(N, K) = f(1/K)$.

Here, $f:(0, 1) \mapsto \mathbb{R}$ *is the function defined by* [2]

$$f(\alpha) = 1 + \left(\frac{1-\alpha}{2}\right)\log\left(\frac{1-\alpha}{2}\right) + \left(\frac{1+\alpha}{2}\right)\log\left(\frac{1+\alpha}{2}\right). \tag{2.8}$$

Notice that the function f defined by (2.8) satisfies $f(\alpha) = 1 - H(\frac{1}{2}(1-\alpha))$, where H is the binary entropy function. In particular, f is positive and continuous for $\alpha \in (0, 1)$. Before going ahead with the main part of the proof, we need to develop some lemmas that will be our main tools.

Lemma 2.1. *Let* $X_1, X_2, \dots, X_N$ *be independent binomial random variables with* $\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$ *for every* i.
(a) *Let* $u_i \in [-1, 1]$, $i = 1, \dots, N$. *Then, for every* $\alpha \in (0, 1)$, *we have*

$$\Pr\left(\frac{1}{N}\sum_{i=1}^{N} u_i X_i \geq \alpha\right) \leq 2^{-Nf(\alpha)}. \tag{2.9}$$

[2] In the definition of f, and throughout the rest of the paper, all logarithms are taken with base 2.

(b)

$$\lim_{N\to\infty}\frac{1}{N}\log \Pr\left(\frac{1}{N}\sum_{i=1}^{N}X_i\geq\alpha\right)=-f(\alpha). \tag{2.10}$$

Proof. Part (b) is obtained from the classical Chernoff bound [1] or from counting arguments [2]. Part (a) also follows from the Chernoff bound, if $u_i=1$ for all i. It remains to prove part (a) for the general case of $u_i\in[-1, 1]$.

We first note that because of the symmetry in the distribution of X_i, we can assume, without any loss of generality that $u_i\in[0, 1]$ for all i. We then have

$$\Pr\left(\frac{1}{N}\sum_{i=1}^{N}u_iX_i\geq\alpha\right)\leq\inf_{s>0}\prod_{i=1}^{N}E[\mathrm{e}^{s(u_iX_i-\alpha)}]\leq\inf_{s>0}\prod_{i=1}^{N}E[\mathrm{e}^{s(X_i-\alpha)}]=2^{-Nf(\alpha)}.$$

The first inequality is obtained by following the steps in the standard proof of the Chernoff bound; the second inequality is obtained by verifying that $\mathrm{e}^{su}+\mathrm{e}^{-su}\leq\mathrm{e}^{s}+\mathrm{e}^{-s}$ for all $u\in[0, 1]$; finally, the final equality is a simple calculation which is also part of the classical proof of the Chernoff bound. □

One consequence of Lemma 2.1 is that for any $\varepsilon>0$, there exists some $N_0(\alpha, \varepsilon)$ such that

$$\Pr\left(\frac{1}{N}\sum_{i=1}^{N}X_i\geq\alpha\right)\geq 2^{-N(f(\alpha)+\varepsilon)},\quad \forall N\geq N_0(\alpha, \varepsilon). \tag{2.11}$$

The following lemma strengthens (2.11) and will be needed later in the proof.

Lemma 2.2. *Let $X_1,\dots,X_N$ be as in Lemma 2.1. Let $\Theta_N=\{(\theta_1,\dots,\theta_N)\in\mathbb{R}^N \mid \sum_{i=1}^{N}|\theta_i|=N\}$. Then, for any $\varepsilon_1>0$, there exists some $N_1(\alpha\varepsilon_1)$ such that*

$$\Pr\left(\frac{1}{N}\sum_{i=1}^{N}\theta_iX_i\geq\alpha\right)\geq 2^{-N(f(\alpha)+\varepsilon_1)},\quad \forall N\geq N_1(\alpha\varepsilon_1),\ \forall\theta\in\Theta_N. \tag{2.12}$$

Proof. Note that the random variables $\sum_{i=1}^{N}\theta_iX_i$ and $\sum_{i=1}^{N}|\theta_i|X_i$ have the same probability distribution. Therefore, without loss of generality, we can and will assume that $\theta_i\geq 0$ for all i. We have

$$\begin{aligned}\Pr\left(\sum_{i=1}^{N}\theta_iX_i\geq\alpha N\right)&=\Pr\left(\sum_{i=1}^{N}\theta_iX_i\geq\alpha N\,\middle|\,\sum_{i=1}^{N}X_i\geq\alpha N\right)\cdot\Pr\left(\sum_{i=1}^{N}X_i\geq\alpha N\right)\\&\geq 2^{-N(f(\alpha)+\varepsilon_1/2)}\Pr\left(\sum_{i=1}^{N}\theta_iX_i\geq\alpha N\,\middle|\,\sum_{i=1}^{N}X_i\geq\alpha N\right),\end{aligned} \tag{2.13}$$

where the last inequality holds for all N large enough, as a consequence of (2.11).

Given any sequence $X=(X_1,\dots,X_N)$, let X^k be its cyclic shift by k positions; that is, $X^k=(X_{k+1}, X_{k+2},\dots,X_N, X_1,\dots,X_k)$. Let X_i^k be the i-th component of X^k. By symmetry, the conditional distribution of X and X^k, conditioned on the event $\sum_{i=1}^{N}X_i\geq\alpha N$, is the same. Therefore,

$$\begin{aligned}\Pr\left(\sum_{i=1}^{N}\theta_iX_i\geq\alpha N\,\middle|\,\sum_{i=1}^{N}X_i\geq\alpha N\right)&=\frac{1}{N}\sum_{k=1}^{N}\Pr\left(\sum_{i=1}^{N}\theta_iX_i^k\geq\alpha N\,\middle|\,\sum_{i=1}^{N}X_i\geq\alpha N\right)\\&\geq\frac{1}{N}\Pr\left(\exists k \text{ such that } \sum_{i=1}^{N}\theta_iX_i^k\geq\alpha N\,\middle|\,\sum_{i=1}^{N}X_i\geq\alpha N\right)\\&=\frac{1}{N}.\end{aligned} \tag{2.14}$$

The last equality follows because if $\sum_{i=1}^{N} X_i \geq \alpha N$, then

$$\sum_{k=1}^{N} \sum_{i=1}^{N} \theta_i X_i^k = \sum_{i=1}^{N} \theta_i \sum_{i=1}^{N} X_i \geq \alpha N^2,$$

which immediately implies that there exists some k for which $\sum_{i=1}^{N} \theta_i X_i^k \geq \alpha N$.

We conclude that (2.13) becomes

$$\Pr\left(\sum_{i=1}^{N} \theta_i X_i \geq \alpha N \right) \geq \frac{1}{N} 2^{-N(f(\alpha)+\varepsilon_1/2)} \geq 2^{-N(f(\alpha)+\varepsilon_1)},$$

where the last inequality follows if N is large enough so that $1/N \geq 2^{-N\varepsilon_1/2}$. □

Having finished with the probabilistic preliminaries, we can now continue with the main part of the proof of Theorem 2.2. We will start with the proof of part (a).

Lemma 2.3. *Suppose that the length n of an input sequence $u \in U_n$ is smaller than $2^{Nf(1/K)-1} - N + 2\lceil N/K \rceil - 1$. Then, there exists some $h \in \{-K\delta/N, K\delta/N\}^N$ such that $\| u * h \|_\infty < \delta$.*

Proof. Let n be as in the statement of the lemma. We will show the existence of such an h by showing that a random element of $\{-K\delta/N, K\delta/N\}^N$ satisfies $\| u * h \|_\infty < \delta$ with positive probability. Indeed, let h be such a random element, under the uniform distribution on $\{-K\delta/N, K\delta/N\}^N$. Then

$$\begin{aligned} \Pr(\| u * h \|_\infty \geq \delta) &\leq \sum_{j=1}^{N+n} \Pr(|(u * h)_j| \geq \delta) = \sum_{j=\lceil N/K \rceil+1}^{N+n-\lceil N/K \rceil+1} \Pr(|(u * h)_j| \geq \delta) \\ &\leq (N + n - 2\lceil N/K \rceil + 1) \max_{1 \leq j \leq N+n} \Pr(|(u * h)_j| \geq \delta), \end{aligned} \tag{2.15}$$

where the equality on the first line holds because for $j \leq \lceil N/K \rceil$, we have

$$|(u * h)_j| = \left| \sum_{i=1}^{N} h_i u_{j-i} \right| = \left| \sum_{i=1}^{j-1} h_i u_{j-i} \right| \leq (j-1) \frac{K\delta}{N} \leq \left(\left\lceil \frac{N}{K} \right\rceil - 1 \right) \frac{K\delta}{N} < \delta$$

and for $j \geq N + n - \lceil N/K \rceil + 2$, we have

$$|(u * h)_j| = \left| \sum_{i=1}^{N} h_i u_{j-i} \right| = \left| \sum_{i=j-n}^{N} h_i u_{j-i} \right| \leq (N - j + n + 1) \frac{K\delta}{N} \leq \left(\left\lceil \frac{N}{K} \right\rceil - 1 \right) \frac{K\delta}{N} < \delta.$$

Furthermore,

$$\begin{aligned} \Pr(|(u * h)_j| \geq \delta) &= \Pr\left(\left| \sum_{i=1}^{N} h_i u_{j-1} \right| \geq \delta \right) \\ &= \Pr\left(\frac{1}{N} \left| \sum_{i=1}^{N} (Nh_i/K\delta) u_{j-i} \right| \geq \frac{1}{K} \right) \leq 2 \cdot 2^{-Nf(1/K)}. \end{aligned} \tag{2.16}$$

The last inequality follows from Lemma 2.1 (a), because the random variables $Nh_i/K\delta$ are independent, take values in $\{-1, 1\}$, and each value is equally likely. Combining (2.15) and (2.16), we conclude that

$$\Pr(\| u * h \|_\infty \geq \delta) \leq 2\left(N + n - 2\left\lceil \frac{N}{K} \right\rceil + 1 \right) 2^{-Nf(1/K)}. \tag{2.17}$$

If $2(N + n - 2\lceil N/K \rceil + 1) < 2^{Nf(1/K)}$, then the right-hand side of (2.17) is smaller than 1. This implies that there exists some $h \in \{-K\delta/N, K\delta/N\}^N$ for which $\| h * u \|_\infty < \delta$. □

Suppose now that the length n of the input sequence u is as in Lemma 2.3, and let the unknown system h have the properties described in that lemma. Since $|(h * u)_i| < \delta$ for all i, there is a choice of the disturbance sequence d that leads to zero output. Consider next the case where the unknown system is actually equal to $-h$. We also have $|(-h * u)| < \delta$, for all i, and a zero output sequence is still possible. Thus, if the output sequence is equal to zero, both h and $-h$ could be the true system. For any identification algorithm, the worst-case error will be at least equal to one half of the distance of these two systems, which is $\|h\|_1 = K\delta$. In fact, the same argument can be carried out if h is replaced by $(1+\varepsilon)h$, where $\varepsilon > 0$ is small enough so that the property $(1+\varepsilon)|(h * u)_i| < \delta$ holds. We can then conclude that the worst-case diameter will be at least $2(1+\varepsilon)K\delta$. We have therefore shown that if $n < 2^{Nf(1/K)-1} - N + 2\lceil N/K \rceil - 1$, then $D_{N,n}(u) > 2K\delta$. Equivalently, $n^*(N, K) \geq 2^{Nf(1/K)-1} - N + 2\lceil N/K \rceil - 1$, which completes the proof of part (a).

We now turn to the proof of part (b) of the theorem. Part (a) implies that $\liminf_{N\to\infty}(1/N) \log n^*(N, K) \geq f(1/K)$. The proof will be completed by showing that

$$\limsup_{N\to\infty} (1/N) \log n^*(N, K) \leq f(1/K).$$

To show this, we have to show the existence of an input sequence u of length close to $2^{Nf(1/K)}$ that results in an uncertainty set of diameter bounded by $2K\delta$. Although we are not able to provide an explicit construction of such an input sequence, we will prove its existence using a probabilistic argument.

We now provide the details of the construction of the input sequence u. Let us fix some $\varphi > 0$. Let $M(N)$ be the smallest integer larger than

$$M(N) \geq 2^{N(f(\varepsilon+1/K)+2\varepsilon)}. \tag{2.18}$$

For every $k \in \{1, \ldots, M(N)\}$, we choose a vector $u^k = (u_1^k, \ldots, u_N^k) \in \{-1, 1\}^N$. The input u is then defined by

$$u = (u^1, u^2, \ldots, u^{M(N)}), \tag{2.19}$$

and has length $NM(N)$.

Lemma 2.4. *Let the input u be constructed as in the preceding paragraph. Furthermore suppose that the entries of the vectors u^k are independent random variables, with each value in the set $\{-1, 1\}$ being equally likely. Then, there exists some $N_2(\varepsilon)$ such that*

$$\Pr(\exists h \in \mathcal{M}_N \text{ such that } \|h\|_1 \geq K\delta,\ \|u * h\|_\infty \leq \delta) < 1, \quad \forall N \geq N_2(\varepsilon). \tag{2.20}$$

Proof. Let Q_N be the left-hand side of (2.20). Notice that if i is an integer multiple of N, with $i = mN$, we have

$$(u * h)_i = \sum_{j=1}^{N} u_j^m h_{N-j}, \quad i = mN. \tag{2.21}$$

We then have

$$\begin{aligned} Q_N &= \Pr(\exists h \in \mathcal{M}_N \text{ such that } \|h\|_1 \geq K\delta,\ \|u * h\|_\infty \leq \delta) \\ &= \Pr(\exists h \in \mathcal{M}_N \text{ such that } \|h\|_1 = K\delta,\ \|u * h\|_\infty \leq \delta) \\ &= \Pr(\exists h \in \mathcal{M}_N \text{ such that } \|h\|_1 = N,\ \|u * h\|_\infty' \leq N/K) \\ &\leq \Pr\left(\exists h \in \mathcal{M}_N \text{ such that } \|h\|_1 = N,\ \left|\sum_{j=1}^{N} u_j^m h_{N-j}\right| \leq N/K,\ m = 1, \ldots, M(N)\right), \end{aligned} \tag{2.22}$$

where the last inequality follows from (2.21).

Let us choose a finite subset $\mathscr{M}_N^\varepsilon$ of $\mathscr{M}_N$ such that for every $h \in \mathscr{M}_N$ with $\|h\|_1 = N$, there exists some $h' \in \mathscr{M}_N^\varepsilon$ satisfying $\|h'\|_1 = N$ and $\|h - h'\|_\infty < \varepsilon$. In particular, $\mathscr{M}_N^\varepsilon$ can be chosen as a subset of the set of all elements of $\mathscr{M}_N$ for which each component is bounded by N and is an integer multiple of ε/N. It is then clear that $\mathscr{M}_N^\varepsilon$ can be assumed to have cardinality bounded by $((2N+1)/\varepsilon)^N$. We then have

$$\begin{aligned}
&\Pr\left(\exists h \in \mathscr{M}_N \text{ such that } \|h\|_1 = N, \left|\sum_{j=1}^{N} u_j^m h_{N-j}\right| \le N/K,\ m = 1,\dots,M(N)\right)\\
&\quad \le \Pr\left(\exists h' \in \mathscr{M}_N^\varepsilon \text{ such that } \left|\sum_{j=1}^{N} u_j^m h'_{N-j}\right| < N(\varepsilon + 1/K),\ m = 1,\dots,M(N)\right)\\
&\quad \le \left(\frac{2N+1}{\varepsilon}\right)^N \max_{h' \in \mathscr{M}_N^\varepsilon} \Pr\left(\left|\sum_{j=1}^{N} u_j^m h'_{N-j}\right| < N(\varepsilon + 1/K),\ m = 1,\dots,M(N)\right).
\end{aligned} \tag{2.23}$$

We provide an upper bound to the probability in the right-hand side of (2.23) by applying Lemma 2.2. (Here, u_j^m and h'_{N-j} correspond to X_i and θ_i in the notation of that lemma.) Indeed, Lemma 2.2 is applicable because $\|h'\|_1 = N$ and the components of the input are i.i.d random variables, with the same distribution as the variables X_i of Lemma 2.1. A minor difference is that the components of h' could be negative, while in Lemma 2.2 we assumed that the components of θ are nonnegative. Nevertheless, if we replace each component of h' with its absolute value, the distribution of the random variables $\sum_{j=1}^N u_j^m h'_{N-j}$ remains the same. We therefore conclude that there exists some $N_2(K, \varepsilon)$ such that

$$\Pr\left(\left|\sum_{j=1}^{N} u_j^m h'_{N-1}\right| < N(\varepsilon + 1/K)\right) \le 1 - 2^{-N(f(\varepsilon+1/K)+\varepsilon)}, \quad \forall m, \forall N \ge N_2(K, \varepsilon). \tag{2.24}$$

By combining (2.22), (2.23), (2.24), and using the statistical independence of the vectors u^m, we obtain

$$\begin{aligned}
Q_N &\le ((2N+1)/\varepsilon)^N \left(1 - 2^{-N(f(\varepsilon+1/K)+\varepsilon)}\right)^{M(N)}\\
&\le ((2N+1)/\varepsilon)^N \exp\{-M(N) 2^{-N(f(\varepsilon+1/K)+\varepsilon)}\} \le ((2N+1)/\varepsilon)^N \exp\{-2^{\varepsilon N}\},
\end{aligned} \tag{2.25}$$

where the second inequality follows from the fact $(1 - 1/x)^x \le e^{-1}$, for every $x > 0$, and the last inequality follows from the definition of $M(N)$ [cf. (2.18)]. It is then easily seen that Q_N converges to zero as N increases, which establishes the desired result. □

Lemma 2.4 establishes that, if the input u is constructed randomly as in the discussion preceding the lemma, then, with positive probability, u will have property P below:

$$\text{P:} \quad \text{if } h \in \mathscr{M}_N \text{ and } \|u * h\|_\infty \le \delta, \text{ then } \|h\|_1 \le K\delta. \tag{2.26}$$

In particular, there exists at least one u, of length $n = M(N)N$ that has property P. [3]

Lemma 2.5. *If an input u has property* P *of* (2.26), *then $D_{N,n}(u) \le 2K\delta$.*

Proof. We apply the input u and measure the output $y = h * u + d$, where h is the unknown plant and d is the disturbance sequence. Given the observed output y, we can infer that h belongs to the set of uncertainty

$$S_{N,n}(y, u) = \{\phi \in \mathscr{M}_N \mid \|y - \phi * u\|_\infty \le \delta\}.$$

Let χ and ψ be two elements of $S_{N,n}(y, u)$. Then, $\|y - \chi * u\|_\infty \le \delta$ and $\|y - \psi * u\|_\infty \le \delta$. Using the triangle inequality, we obtain $\|u * (\chi - \psi)/2\|_\infty \le \delta$. Since u has property P, we conclude that

[3] In fact, it is easily seen that Q_N converges to zero very rapidly, which implies that most u's will have property P.

$\|(\chi-\psi)/2\|_1 \le K\delta$ or $\|\chi-\psi\|_1 \le 2K\delta$. Since this is true for all elements of $S_{N,n}(y, u)$, the diameter of $S_{N,n}(y, u)$ is at most $2K\delta$. □

As discussed earlier, if N is large enough, there exists an input of length $n = M(N)N$ that has property P and, by Lemma 2.5, leads to uncertainty sets whose diameter is bounded above by $2K\delta$. It follows that $n^*(N, K) \le M(N)N$. Using the definition of $M(N)$ [cf. (2.18)], we see that

$$\limsup_{N\to\infty}(1/N)\log n^*(N, K) \le \limsup_{N\to\infty}(1/N)\log M(N)N \le f\left(\varepsilon + \frac{1}{K}\right) + 2\varepsilon. \tag{2.27}$$

Since Eq. (2.27) is valid for all $\varepsilon > 0$, and since f is continuous, we conclude that

$$\limsup_{N\to\infty}(1/N)\log n^*(N, K) \le f(1/K),$$

which concludes the proof of Theorem 2.2. □

3. Conclusions

This paper addresses issues in the sample complexity of worst-case identification in the presence of unknown but bounded noise. Two main results are furnished: the first is a lower bound on the length of inputs necessary to approximate N steps of an impulse response to an accuracy within a factor K of the best possible achievable error. This bound has the form $2^{Nf(1/K)}$, and hence is exponential in N. The second result shows that this lower bound in asymptotically tight, i.e. for large enough N, there exists an input of length close to the lower bound that allows the identification of N steps of the impulse response.

References

[1] R.R. Bahadur, *Some Limit Theorems in Statistics* (SIAM, Philadelphia, PA, 1971).
[2] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Channels* (Academic Press, New York, 1981).
[3] M.A. Dahleh, E. Sontag, D.N. Tse and J.N. Tsitsiklis, Worst-case identification for a class of nonlinear fading memory systems, in *Proc. American Control Conference* (1992).
[4] G. Gu, P.P. Khargonekar and Y. Li, Robust convergence of two stage nonlinear algorithms for identification in H_∞, *Systems Control Lett.* **18** (4) (1992) 253–263.
[5] G. Gu and P.P. Khargonekar, Linear and nonlinear algorithms for identification in $\mathscr{H}_\infty$ with error bounds, *IEEE Trans. Automat. Control* **37** (7) (1992) 953–963.
[6] A.J. Helmicki, C.A. Jacobson and C.N. Nett, Control-oriented system identification: A worst-case/deterministic approach in H_∞, *IEEE Trans. Automat. Control* **36** (10) (1991) 1163–1176.
[7] A.J. Helmicki, C.A. Jacobson and C.N. Nett, Identification in H_∞: Linear algorithms, *Proc. American Control Conference* (1990) 2418–2423.
[8] C.A. Jacobson and C.N. Nett, Worst-case system identification in ℓ_1: Optimal algorithms and error bounds, *Proc. American Control Conference* (June 1991).
[9] B. Kacewicz and M. Milanese, On the optimal experiment design in the worst-case ℓ_1 system identification, *Proc. Conf. Dec. Control* (1992) 56–61.
[10] P.M. Makila, Robust Identification and Galois Sequences, Technical Report 91-1, Process Control Laboratory, Swedish University of Abo (Jan. 1991).
[11] P.M. Makila and J.R. Partington, Robust approximation and identification in H_∞, *Proc. American Control Conference* (June 1991).
[12] M. Milanese and R. Tempo, Optimal algorithm theory for robust estimation and prediction, *IEEE Trans. Automat. Control* **30** (1985) 730–738.
[13] M. Milanese, Estimation theory and prediction in the presence of unknown and bounded uncertainty: A survey, in: M. Milanese, R. Tempo and A. Vicino, Eds., *Robustness in Identification and Control* (Plenum Press, New York, 1989).
[14] K. Poolla and A. Tikku, On the time complexity of worst-case system identification, Preprint (1992).

[15] J.F. Traub and H. Wozniakowski. *A General Theory of Optimal Algorithms* (Academic Press. New York. 1980).
[16] J.F. Traub, G. Wasilkowski and H. Wozniakowski. *Information-Based Complexity* (Academic Press. New York, 1988).
[17] D. Tse, M.A. Dahleh and J.N. Tsitsiklis. Optimal asymptotic identification under bounded disturbances. *IEEE Trans. Automat. Control* (to appear).
[18] D.N.C. Tse, M.A. Dahleh. J.N. Tsitsiklis. Optimal and robust identification in the ℓ_1 norm. *Proc. American Control Conference* (June 1991).
[19] G. Zames. On the metric complexity of causal linear systems: ε-entropy and ε-dimension for continuous time. *IEEE Trans. Automat. Control* **24** (2) (1979) 222–230.

A Framework for Iterative Modeling and Control

Mitchell M. Livstone and Munther A. Dahleh
Electrical Engineering and Computer Science
Massachusetts Institute of Technology

Abstract

In this paper we present a new framework for iterative modeling and control. We begin by describing the unknown process with an uncertain model whose parametrization depends on prior information, available control design tools and other modeling preferences. This is formally presented as a model set transformation problem. The second step is an iterative procedure for refining the uncertainty set via robust control based model invalidation and can be viewed as a systematic way of efficiently searching for a controller delivering a certain desired level of performance to the unknown process. As a result, either the performance goal will be met or the entire uncertainty set will be invalidated in accordance with our modeling and control method prejudice. An iterative scheme based on a special model structure and rank one mixed μ synthesis will be described in detail and a specific example will be used to illustrate the ideas.

1 Introduction

Over the past decade, there has been much research activity in the area of worst–case, or control–oriented system identification. The motivation can be attributed to new advances in robust control theory which did not interface well with existing theory of classical system identification. The main focus of this research has been the design of algorithms that yield nominal models along with measures of uncertainty which are well suited for robust control design [11, 10, 28, 20, 14]. Unfortunately, these worst–case algorithms tend to provide error bounds which are very conservative in practice [13] and are therefore of limited utility. This is one motivation for the area of iterative identification and control which has recently gained attention in the control

community. Several researchers have been working on the "connections" between identification and control [1, 30, 31, 9, 17, 24, 25, 16, 15]. This work has typically been in the spirit of adaptive control. In other words, a sequence of nominal models is being identified, while a sequence of corresponding robust controllers is being designed for these models. The hope is that the models are, in some sense, getting closer to the unknown process and the performance is improving.

The more recent formulation of Dahleh and Doyle [4] uses an entirely different philosophy. Here, the goal is to observe some experimentally generated finite set of data and find a controller that meets a given performance specification for the unknown process. The model is thought of as a tool which is chosen based on the designer's preferences of control design techniques and ways of explaining the observed data. In this sense, the chosen model parameterization is good only if a controller designed for this model can also achieve good performance with the unknown process. On the other hand, if a controller delivers good performance with the model, yet fails to meet the performance specifications with the unknown process, this model is considered to be a poor description of the process and should be invalidated. In this way, a conservative model set can be effectively shrunk until the remaining elements can deliver a controller which will achieve the desired performance specifications on the actual process, or the whole set is invalidated.

In this paper we develop this philosophy further and give a concrete example of an iterative scheme based on such model invalidation through robust control design. The next section considers the modeling step at a general level. Section 3 describes the general philosophy of an iterative scheme and comments on the computation issues in the general case. This is followed by the development of an iterative scheme based on a fixed pole model (FPM) and the rank one mixed μ synthesis (ROS) robust control technique. Section 4 discusses the modeling step in more detail and presents some new results in this direction. This is followed by a detailed descripion of an iterative scheme based on the FPM and ROS. The computations associated with the ROS scheme are outlined in Section 5.1 and some worst-case complexity results are also collected there. Finally, a specific example of the ROS based iterative scheme is considered in Section 7.

2 Step One: Selection of Model Parameterization

In general, the selection of the model parameterization is a process that requires engineering insight as well as careful consideration of available robust control techniques and available information about the process to be controlled. Currently there are few robust control methods available and the existing methods incorporate only special types of performance objectives and uncertainty structures [6, 5, 29, 23, 7, 12]. All of these design methods can accommodate unmodeled dynamics as uncertainty and norms of weighted transfer functions as design specifications. The mixed–μ (rank one) synthesis method of Rantzer and Megretskii [23] can also nonconservatively accommodate parametric uncertainty models (having some structural restrictions). Once the desired model structure, or parameterization is chosen, the problem is to efficiently map the prior information about the unknown process into a model set having the desired structure. This step can be extremely difficult and so the structure of the prior information can significantly influence the choice of model parameterization.

This model set transformation step can be made more rigorous by defining the original prior information set to be the set of models consistent with the priors

$$\mathcal{M}_{\text{prior}} = \{P(a) : a \in A_p \subset \mathbb{R}^n\} \tag{1}$$

where the prior information is imbedded in A_p and the functional dependence of $P(a)$ on a. Note that any unmodeled dynamics are also contained in $\mathcal{M}_{\text{prior}}$. Given the desired model structure parameterization, $G(\theta, \Delta)$, the goal is to find the smallest $\epsilon \geq 0$ and $\Theta_o \subset \mathbb{R}^m$ such that

$$\mathcal{M}_{\text{prior}} \subseteq \mathcal{M}_{\text{des}} \equiv \{G(\theta, \Delta) : \theta \in \Theta_o, \|\Delta\|_\infty \leq \epsilon\}$$

This is generally a very difficult problem to solve and an approximate solution for the special FPM case will be developed later in Section 4.

3 Step Two: Inner Loop

The main objective of an iterative scheme in our framework is to efficiently search over the auxilliary (model) space for a controller which meets the given performance objective with the actual unknown process. The idea is to partition the model set and then efficiently invalidate subsets in the partition

while searching for a subset which yields a controller that meets the desired performance objective for the actual process. A subset is invalidated if there exists a controller which achieves a certain level of robust performance for this set but fails to achieve the same level of performance for the actual process. The key issues are choosing the model structure/parameterization and deciding how to partition the model set. Clearly, the termination of such an iterative scheme can occur only if performance can be tested using a finite duration signal. We now make these ideas more rigorous before stating the iterative procedure.

3.1 Preliminaries

We first establish some useful notation. The model set $(\mathcal{M}_0, \mathcal{D})$ is assumed to describe the unknown process P_r. The input/output relations for P_r and some plant/noise pair $(h, d) \in (\mathcal{M}_0, \mathcal{D})$ are written as $y = P_r u$ and $y = (h, d)u$, respectively. The reason for using this notation is that the development remains general, and we are not forced to write $y = h * u + d$ or $y = h * (u + d)$, for example. Let the process augmented by the exogenous inputs w and measured outputs z be denoted by P. Now, assume that some controller K is used to close the loop around the augmented process. The performance objective is to minimize (in some appropriate sense) the measured output z for some exogenous input w.

We assume that w and z lie in some spaces of sequences of p and q dimensional vectors. The relation for w and z is now some LFT of the process and controller and will be represented as $\mathcal{F}_l(P, K)$. With a slight abuse of notation we will similarly represent the closed loop relation with (h, d) in place of P as $\mathcal{F}_l((h, d), K)$ and the set of relations with $(\mathcal{M}_0, \mathcal{D})$ as $\mathcal{F}_l((\mathcal{M}_0, \mathcal{D}), K)$. Finally, given an exogenous signal w_0, assuming uniqueness of solutions in the closed loop system, we can define a map G which takes (w_0, P, K) into u_0, the input to the plant, which we write as $u_0 = G(w_0, P, K)$. We next establish the way in which we view performance.

When an engineer designs a control system to achieve certain performance objectives, the design is tested and will seldom give satisfactory performance on the first try. The main point is that during the testing phase, only a finite time experiment is available. This means that the engineer can increase his/her confidence by observing the closed loop system for some finite set of exogenous inputs thought to represent the *typical* signals which the system will have to face in the future. Following this philosophy, we view performance in relationship to some *finite collection of finite duration*

exogenous signals and the metrics used on the input and output spaces.

One can think of the standard model validation problem in the following way. Given $(\mathcal{M}_0, \mathcal{D})$, u and $y = P_r u$, if $(\mathcal{M}_0, \mathcal{D})$ is inconsistent with (u, y), then we say that $(\mathcal{M}_0, \mathcal{D})$ does not adequately describe P_r *with respect to the input* u. The computations associated with this consistency test is the main focus of the recent work on model validation [26, 27, 22]. We now show that performing model invalidation based on observed performance (i.e., control based model invalidation) is related to standard model invalidation using a special input, u^*, which is well defined.

Assume that a set of exogenous inputs, $\mathcal{W} = \left\{ w_i \in R^{N \times p} \mid i \in [1, M_w] \right\}$, is given and the goal is to achieve $\|\mathcal{F}_l(P, K) w_i\| \leq \gamma \|w_i\| \;\; \forall i \in [1, M_w]$. The choice of norms here dictates which control design methodology one will have to utilize [5, 19]. We propose the following alternative way of thinking about model validation, which is similar to the idea in [4].

Definition 3.1 *Given a model set* $(\mathcal{M}_0, \mathcal{D})$ *and a* K *which achieves*
$\|\mathcal{F}_l((\mathcal{M}_0, \mathcal{D}), K) w_i\| \leq \gamma \|w_i\| \;\; \forall i \in [1, M_w]$,
if $\exists j \in [1, M_w]$ *s.t.* $\|\mathcal{F}_l(P, K) w_j\| > \gamma \|w_j\|$,
then *we say that* $(\mathcal{M}_0, \mathcal{D})$ *does not adequately describe* P *in view of the objectives.*

In other words, if we can find a controller K which achieves the desired performance for $(\mathcal{M}_0, \mathcal{D})$, yet does not achieve it for the process P_r, we invalidate $(\mathcal{M}_0, \mathcal{D})$. The following result establishes a connection between this viewpoint and the standard model validation problem.

Lemma 3.2 *Given* $(\mathcal{M}_0, \mathcal{D})$ *and* $\mathcal{W}$, *let* K *be a controller which achieves*

$$\|\mathcal{F}_l((\mathcal{M}_0, \mathcal{D}), K) w_i\| \leq \gamma \|w_i\| \;\; \forall i \in [1, M_w]$$

Furthermore, for each $i \in [1, M_w]$, *let* $u_i^* = G(w_i, P, K)$ *and* $y_i^* = P u_i^*$.
If $\exists \{(h_i, d_i)\}_1^{M_w} \subset (\mathcal{M}_0, \mathcal{D})$ *s.t.* $y_i^* = (h_i, d_i) u_i^* \; \forall i \in [1, M_w]$
then $\|\mathcal{F}_l(P, K) w_i\| \leq \gamma \|w_i\| \; \forall i \in [1, M_w]$.

Proof. Given $\mathcal{W}$, take any $i \in [1, M_w]$. During the experiment, the closed loop system generates unique u_i^*, y_i^* and $z_i^* = \mathcal{F}_l(P, K) w_i$. By assumption, $y_i^* = (h_i, d_i) u_i^*$ (i.e., (h_i, d_i) can interpolate u_i^* to y_i^*). Because of uniqueness of solutions, if we consider (h_i, d_i) in the loop instead of P, the resulting z must be the same as z_i^* from the experiment. This means exactly that

$$\mathcal{F}_l(P, K) w_i = \mathcal{F}_l((h_i, d_i), K) w_i \tag{2}$$

and since K was designed to achieve $\|\mathcal{F}_l((\mathcal{M}_0, \mathcal{D}), K)w_i\| \leq \gamma\|w_i\|$, it certainly achieves $\|\mathcal{F}_l((h_i, d_i), K)w_i\| \leq \gamma\|w_i\|$. In view of equation 2, it is clear that this implies $\|\mathcal{F}_l(P, K)w_i\| \leq \gamma\|w_i\|$. □

A picture showing how these variables are related in the closed loop is shown in Figure 1.

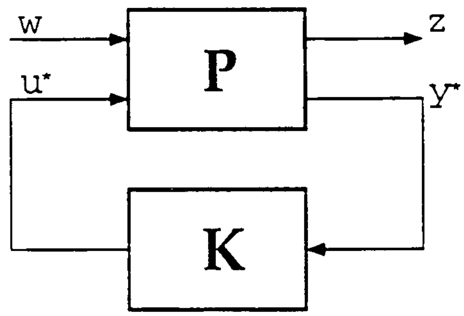


Figure 1: Closed–Loop System Variables Depending on w

This lemma says that if we cannot invalidate the set $(\mathcal{M}_0, \mathcal{D})$ w.r.t. (u_i^*, y_i^*) for any $i \in [1, M_w]$, then we will not be able to invalidate it based on observing the performance w.r.t. $\mathcal{W}$. Notice that the reverse implication is not necessarily true because K is not designed to achieve performance **only** for $(\mathcal{M}_0, \mathcal{D})$. Even if $(\mathcal{M}_0, \mathcal{D})$ is not consistent with (u_i^*, y_i^*), K may inadvertently achieve the desired performance for some $(h, d) \notin (\mathcal{M}_0, \mathcal{D})$ which is consistent with (u_i^*, y_i^*).

The above lemma can be easily extended to performance objectives such as $\|W_p\mathcal{F}_l(P, K)w_i\| \leq \gamma\|w_i\|$, where W_p is some weighting function. In fact, it is easy to see that the result holds for any performance objective that is implied by the robust control design method. In the iterative scheme, the converse of this lemma is actually used and is stated here as a corollary.

Corollary 3.3 *Given the assumptions of Lemma 3.2,* **if** $\exists j \in [1, M_w]$ *s.t.* $\|\mathcal{F}_l(P, K)w_j\| > \gamma\|w_j\|$, **then** *there does not exist any* $(h, d) \in (\mathcal{M}_0, \mathcal{D})$ *s.t.* $y_j^* = (h, d)u_j^*$.

This says that if we invalidate the set $(\mathcal{M}_0, \mathcal{D})$ based on performance, then we would have also invalidated it by using the resulting (u^*, y^*). It is important to reiterate the significance of the unidirectional implication in Lemma 3.2 and Corollary 3.3. This means that even though the observed performance is satisfied, (u^*, y^*) still may not be consistent with $(\mathcal{M}_0, \mathcal{D})$. In this

way, the performance test is only sufficient for invalidation of $(\mathcal{M}_0, \mathcal{D})$ w.r.t. (u^*, y^*). This is not a deficiency in the procedure, but merely reflects the fact that model invalidation in the sense of Definition 3.1 is only sufficient for model invalidation in the traditional sense. In other words, as long as the controller we have designed for $(\mathcal{M}_0, \mathcal{D})$ is delivering the desired performance for the process P, we have no reason to invalidate the set $(\mathcal{M}_0, \mathcal{D})$.

3.2 Inner Loop and Computations

The following two additional assumption are needed since a nonconservative robust control technique does not exist for every type of performance objective.

Assumption 1 There exists a robust control technique which implies the performance objective above (possibly conservative).

Assumption 2 We accept this robust control technique in the sense that if it cannot come up with a controller satisfying a given performance objective, we assume that no controller can satisfy it. We are now in a position to describe the inner loop of the iterative scheme which we show in the flow chart in Figure 2.

The computational difficulty is imbedded in the robust control design step while the efficiency is controlled by the partition of the model set and selection of the candidate subsets. Given the desired performance objective, it may be difficult to determine how small a subset of the partition should be for the robust control problem to have a feasible solution. This means that a practical scheme may be based on further refining the subsets until the robust control problem is solved and only then trying these controllers on the actual process. This is exactly what is done in the iterative scheme described in Section 6.

This general description allows for any model structure and corresponding partition of the model set to be used. In accordance with the above additional assumptions, we only require a robust control technique which implies the given performance specification type. If the robust control design method is conservative, this will be reflected in the conservatism of the iterative scheme in which it is used. Finally, it is important that a finite partition is used to insure termination in finite time. If, for example, the performance objective is too difficult and cannot be met for any plant in the model set, one does not want to refine the partition ad infinitum. There should be some chosen partition level at which a subset giving no feasible robust control solution will be invalidated.

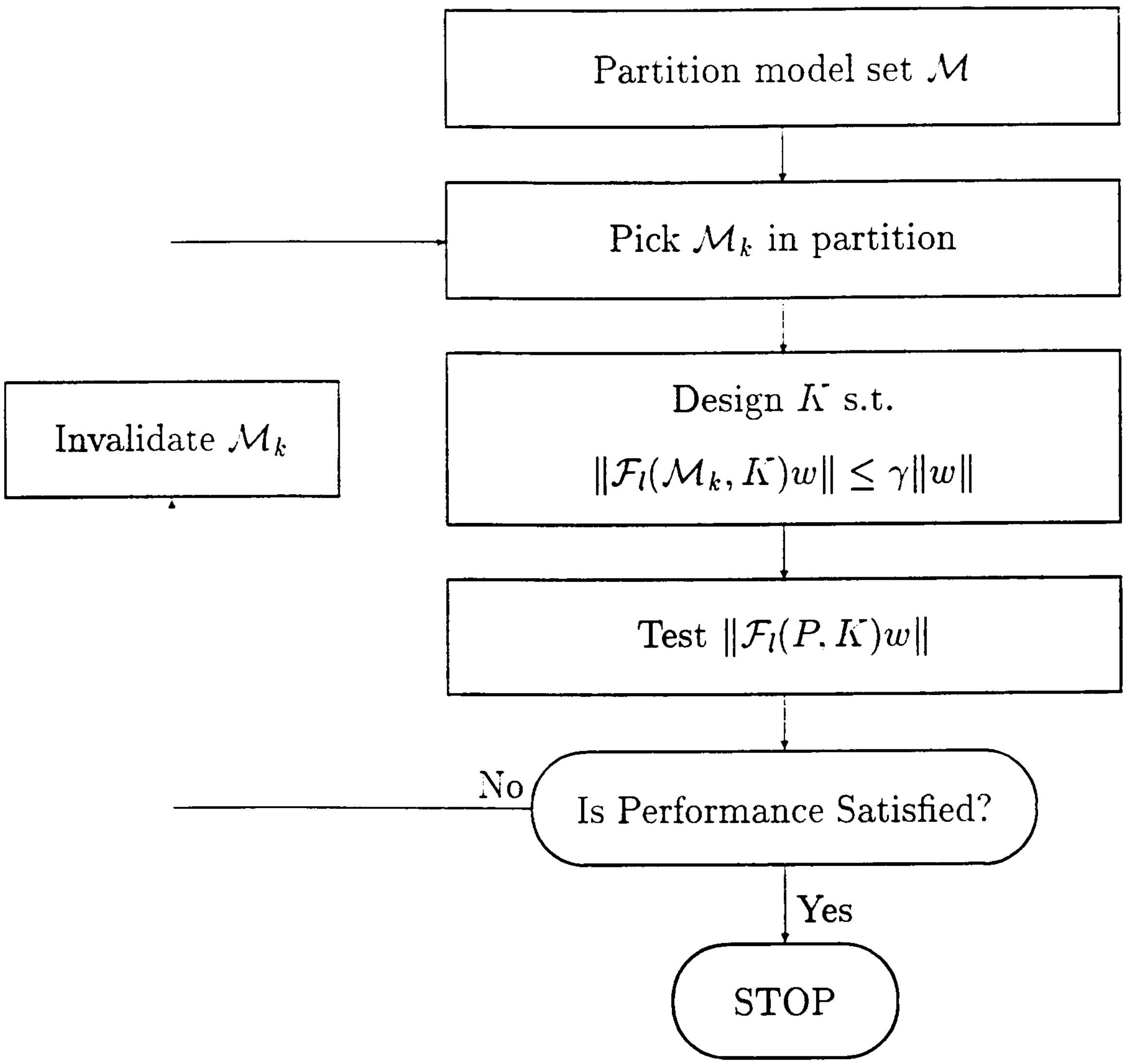


Figure 2: Iterative Procedure

4 Fixed Pole Model for Rank One Synthesis

The ROS design method was developed by Rantzer and Megretskii in [23]. The method is limited to special model structures (i.e., SISO, MISO and SIMO with real and complex coprime factor perturbations), however, the solution is a convex optimization problem. If we also require robust performance such as minimizing a weighted sensitivity (i.e., $\|W_pS\|$), the uncertainty can only be in the *numerator* of the model if the rank one structure is to be maintained. This will be described in more detail in Section 5.1. The corresponding model will be referred to as the fixed pole model (FPM) and is given by

$$G(\theta, \Delta) = \frac{\sum_{k=0}^{m-1} \theta_k z^k + W\Delta}{A(z)} \tag{3}$$

where $\|\Delta\|_\infty \leq 1$, $\theta \in \Theta \subset \Re^m$, W is a stable and invertible weighting function, and $A(z)$ is a stable polynomial.

When the prior information is given in terms of uncertain pole locations or other structures which are not compatible with fixed poles, the mapping of these priors into the appropriate parameter set Θ and a weighting function W can be difficult. In particular, when $W = \epsilon W_o$, one would like to compute the smallest $\epsilon > 0$ and the corresponding set Θ such that the FPM set contains all the plants given by the prior information. It can easily be shown that computing ϵ is equivalent to computing the n–width of the prior model set [18], and finding the corresponding parameter uncertainty set Θ is also a difficult task. For this model and the iterative scheme described in Section 6, it is important to choose W with very little conservatism (i.e., try to lump as much of the pole mismatch into the parametric uncertainty as possible). This is true because the iterative scheme will reduce uncertainty in the parametric part, while the $W\Delta$ part will remain fixed. We next discuss the modeling step for the FPM in more detail.

4.1 Approximate Solution to Model Set Transformation: FPM Case

We assume that the prior model set is stable and of the form

$$\mathcal{M}_{\text{prior}} = \{P(a) + W\Delta : a \in A_p, \|\Delta\|_\infty \leq 1\}$$

where $P(a) = B(z)/(z^n + a_1 z^{n-1} + \cdots + a_n)$, but it is sufficient to consider just the parametric part $P(a)$ since any additive unmodeled dynamics can

be added to the FPM model set. Since we do not know how to solve the exact n-width problem by finding the optimal fixed pole locations for $A(z)$, we use the following approach. Let $A_c(z)$ be the denominator polynomial of $P(a_c)$ where a_c is the center of the set A_p. We now use

$$\frac{\sum_{k=0}^{m-1} \theta_k z^k}{(A_c(z))^{\lfloor m/n \rfloor}}$$

to represent the actual plant. This means that we are using basis functions of the form $\{\frac{z^{nk+j}}{A_c^k(z)}\}$. We now define $A(z) = (A_c(z))^{\lfloor m/n \rfloor}$ and $H_k(z) = z^k/A(z)$. This means that the new model parameterization is given by the subspace $\{\theta^T H : \theta \in \Xi^m\}$. This results in the simplified problem

$$\epsilon^* = \sup_{a \in A_p} \inf_{\theta \in \Xi^m} \|P(a) - \theta^T H\|_\infty \tag{4}$$

This problem is equivalent to finding the maximum deviation from the set $\mathcal{M}_{\text{prior}}$ to a fixed finite dimensional subspace given by $H^T\theta$. We will need the following result for a simplification of the above problem.

Let the state space representation for $\theta^T H$ be

$$\theta^T H(z) \sim \left[\begin{array}{c|c} A_\theta & B_\theta \\ \hline C_\theta & D_\theta \end{array}\right]$$

We can assume WLOG that $D_\theta = 0$ since the constant term can always be matched exactly. Furthermore, in this case we can write

$$\theta^T H = \frac{\sum_{k=0}^{m-1} \theta_k z^k}{z^m + \sum_{k=0}^{m-1} h_k z^k}$$

and then simply take

$$A_\theta = \begin{bmatrix} 0 & 1 & 0 & \cdots & 1 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 \\ -h_0 & -h_1 & \cdots & -h_{m-2} & -h_{m-1} \end{bmatrix} \quad B_\theta = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

and $C_\theta = \theta^T$. Similarly, for a given $a \in A_p$, let the plant $P(a)$ (modulo the DC value) in the prior model set have a representation

$$\left[\begin{array}{c|c} A_P(a) & B_P(a) \\ \hline C_P(a) & 0 \end{array}\right]$$

This means that the system $P(a) - \theta^T H$ has the state space representation

$$\left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \quad A = \begin{bmatrix} A_\theta & 0 \\ 0 & A_{P(a)} \end{bmatrix}$$

$$B = \begin{bmatrix} B_\theta \\ B_{P(a)} \end{bmatrix} \quad C = \begin{bmatrix} -\theta^T & C_{P(a)} \end{bmatrix}$$

Lemma 4.1 *Given the definitions above, for any $a \in A_p$,*

$$\inf_{\theta \in \mathbb{R}^m} \|P(a) - H^T\theta\|_\infty = \inf\{\gamma, \theta, X : (\star)\}$$

where the condition $(\star)$ is equivalent to

$$\exists X > 0 \; s.t. \; \left(\begin{array}{cc} -\begin{bmatrix} A^TXA - X & A^TXB + C^TD \\ B^TXA + D^TC & B^TXB + D^TD - \gamma^2 I \end{bmatrix} & \begin{bmatrix} C^T \\ 0 \end{bmatrix} \\ \begin{bmatrix} C & 0 \end{bmatrix} & I \end{array} \right) \geq 0$$

Proof. The system $P(a) - \theta^T H$ has a state space representation with C affine in θ. It follows almost directly from a theorem in Zhou, et.al. [32] that

$$\|P(a) - \theta^T H\|_\infty < \gamma \text{ if, and only if}$$

$$\exists X > 0 \text{ s.t. } \begin{bmatrix} A^TXA + C^TC - X & A^TXB + C^TD \\ B^TXA + D^TC & B^TXB + D^TD - \gamma^2 I \end{bmatrix} < 0$$

This is an LMI in X and γ^2, but not in θ since this matrix inequality is quadratic in C, and C is an affine function of θ. The result follows after performing a similar trick to the one used in Boyd, et.al. [3]. □

Note that this is now an LMI in the variables γ, θ and X, and can be readily solved using various efficient interior point algorithms given in [21].

We can now present an algorithm for computing ϵ^* and the corresponding set Θ_0 such that

$$\mathcal{M}_{\textbf{prior}} \subseteq \mathcal{M}_{\textbf{des}} \equiv \left\{\theta^T H + \epsilon^* \Delta : \theta \in \Theta_0, \|\Delta\|_\infty \leq 1\right\}$$

The main idea is to relax the problem to one of finding ϵ^* within a small constant $0 < \eta$ which is chosen *a priori*. Because of certain continuity properties, this η determines an ϵ–net for the set A. An LMI is then solved

for each lattice point and yields a finite set of optimal θ's and ϵ's. In the end, ϵ^* is bounded by $\eta +$ the maximum over the computed ϵ's, and the set Θ_0 can be taken as any set which contains the finite set of all the optimal θ's. We now make this more rigorous.

Assume that A_p is compact, $P(a)$ is stable for all $a \in A_p$, and there exists a finite constant, M, such that

$$\sup_{a \in A_p} \|P(a)\|_\infty \leq M < \infty .$$

Let a enter affinely into the denominator of $P(a)$. We will show a bit later that the function mapping $a \longmapsto \|P(a)\|_\infty$ is uniformly continuous on A. This statement is equivalent to saying that for any $\eta > 0$ there exists a δ^* such that for all $a_1, a_2 \in A_p$

$$\|P_{a_1} - P_{a_2}\|_\infty < \eta \text{ whenever } \|a_1 - a_2\| < \delta^*$$

This suggests the following algorithm for computing ϵ^* and Θ_0.

1. Choose some $0 < \eta \ll 1$ and compute the corresponding δ^* (the computation will be discussed shortly).

2. Set up a lattice $\{a_j\}_{j=1}^N$ such that the union of the δ^*–neighborhoods centered at the lattice points is a finite cover for the set A.

$$A \subseteq \bigcup_{j=1}^{N} B_{\delta^*}(a_j)$$

3. For each $j \in [1, N]$, solve (via an LMI)

$$\epsilon_j = \inf_{\theta \in \mathbb{R}^m} \|P(a_j) - \theta^T H\|_\infty \text{ and } \theta_j = \arg\min \|P(a_j) - \theta^T H\|_\infty$$

   and record the pairs $\{(\epsilon_j, \theta_j)\}$.

4. Take $\epsilon^* = \eta + \max_j\{\epsilon_j\}$ and Θ_0 as the smallest hypercube containing $\{\theta_j\}_{j=1}^N$.

This has the following properties.

Theorem 4.2 *Given the above algorithm define* $\bar{\epsilon} \equiv \max_j \epsilon_j$. *Then*

$$\bar{\epsilon} \leq \sup_{a \in A_p} \inf_{\theta \in \mathbb{R}^m} \|P(a) - \theta^T H\|_\infty \leq \bar{\epsilon} + \eta$$

and

$$\{P(a) : a \in A_p\} \subseteq \left\{\theta^T H + \Delta : \theta \in \{\theta_j\}, \|\Delta\|_\infty \leq \bar{\epsilon} + \eta\right\}$$

Proof. The lower bound in the first part is immediate. The upper bound is proven by defining a_* to be the maximizing point in A_p. But then there exists a $k \in [1, N]$ such that $\|a_k - a_*\| \leq \delta^*$ which means that $\|P(a_k) - P(a_*)\|_\infty < \eta$. It is easy to see that f r any $\theta \in \mathbb{R}^{\mho}$ we have

$$\begin{aligned} \|P(a_*) - \theta^T H\|_\infty &= \|P(a_*) - \theta^T H + P(a_k) - P(a_*)\|_\infty \\ &\leq \|P(a_k) - \theta^T H\|_\infty + \|P(a_k) - P(a_*)\|_\infty \\ &\leq \epsilon_k + \eta \leq \bar{\epsilon} + \eta \end{aligned}$$

The second part of the theorem is immediate from the above argument. □

We now show that the function $a \longmapsto \|P(a)\|_\infty$ is indeed uniformly continuous on A_p and in the process, show the steps necessary to compute δ^*. We state the main result in the following theorem.

Theorem 4.3 *Let $A_p \subset \mathbb{R}^n$ be compact and the set $\{P(a) : a \in A_p\}$ is a stable subset of H_∞. Assume that $P(a) = \frac{N}{D_a}$, where D_a is affine in a. Then the map $a \longmapsto \|P(a)\|_\infty$ is uniformly continuous on A_p, meaning*

$\forall \eta > 0, p \in [1, \infty], \exists \delta^*$ *s.t.* $\forall a', a'' \in A_p$, $\|a' - a''\|_p < \delta^* \Rightarrow \|P_{a'} - P_{a''}\|_\infty < \eta$

Furthermore, δ^* satisfies

$$\delta^* \leq \frac{\eta}{n^{1/q} M_T}, \quad 1 = \frac{1}{p} + \frac{1}{q}$$

and the computation of M_T requires either 2 mixed μ analyses or, with a bit more conservatism, 1 mixed μ analysis and 1 LMI solution.

Proof. We first define

$$D_a(e^{i\omega}) \equiv 1 + a^T \xi(\omega) \text{ where } \xi(\omega) = [e^{i\omega} \ \cdots \ e^{in\omega}]$$

and collect a few results. First, because of stability of the set $\{P(a) : a \in A_p\}$, we can compute an upper bound

$$\sup_{a \in A_p} \sup_{\omega \in [0, 2\pi]} \frac{1}{|1 + a^T \xi(\omega)|} < M_D < \infty \tag{5}$$

and use this to get

$$\inf_{a \in A} \inf_{\omega \in [0, 2\pi]} |1 + a^T \xi(\omega)| = \frac{1}{\sup_{a \in A} \sup_{\omega \in [0, 2\pi]} |1 + a^T \xi(\omega)|^{-1}} > \frac{1}{M_D} > 0 .$$

As shown in [18], the computation in Equation 5 requires the solution of one mixed μ analysis problem. Next, using the same argument, we can compute $\sup_{a\in A_p} \|P(a)\|_\infty$ via another mixed μ analysis or by an additional LMI as follows.

$$\sup_{a\in A_p} \|P(a)\|_\infty \leq \|N\|_\infty \left\| \frac{1}{D_a} \right\|_\infty$$

The second quantity was computed in Equation 5 while $\|N\|_\infty$ can be computed via a simple LMI problem. Whichever solution is used, let the computed upper bound for $\sup_{a\in A_p} \|P(a)\|_\infty$ be denoted by M_P. We next define the function

$$g(a,\omega) \equiv \frac{1}{|1 + a^T\xi(\omega)|^2} = \frac{1}{1 + 2a^T Re\{\xi(\omega)\} + a^T\xi(\omega)\xi^*(\omega)a} \tag{6}$$

which we can now differentiate w.r.t. a_k for all $k \in [1, n]$ to get

$$\frac{\partial g(a,\omega)}{\partial a_k} = -\frac{2\cos(k\omega) + 2Re\{e^{-ik\omega}a^T\xi(\omega)\}}{(1 + 2a^T Re\{\xi(\omega)\} + a^T\xi(\omega)\xi^*(\omega)a)^2} \tag{7}$$

We can use implicit differentiation to show that

$$\frac{\partial \left|\frac{1}{1+a^T\xi(\omega)}\right|}{\partial a_k} = \frac{|1 + a^T\xi(\omega)|}{2} \frac{\partial g(a,\omega)}{\partial a_k}$$

and combine everything above to get

$$\begin{aligned}
\left|\frac{\partial |P(a)(e^{i\omega})|}{\partial z_k}\right| &= \frac{|N(e^{i\omega})||1 + a^T\xi(\omega)|}{2} \left|\frac{\partial g(a,\omega)}{\partial a_k}\right| \\
&= \left|\frac{N(e^{i\omega})}{1 + a^T\xi(\omega)}\right| \frac{|\cos(k\omega) + Re\{e^{-ik\omega}a^T\xi(\omega)\}|}{|1 + a^T\xi(\omega)|^2} \\
&\leq\ \leq M_P M_D^2(1 + \|a\|_p n^{1/q}) \\
&\leq M_P M_D^2(1 + n^{1/q} M_A) \leq M_T
\end{aligned}$$

where $M_A = \sup_{a\in A} \|a\|_p$ is bounded since A is compact.

We can use this bound as follows. First, recall the multivariable mean value theorem. If a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ has bounded derivatives, then for any $x, y \in \mathbb{R}^n$ we have

$$f(y) - f(x) = \nabla_{\tilde{x}} f \cdot (y - \tilde{x})$$

for some $\tilde{x} = tx + (1-t)y$, $t \in [0,1]$ (i.e., $\|y - \tilde{x}\| \le \|y - x\|$). This means that if $|\partial f/\partial x_k| \le L$ for all $k \in [1, n]$,

$$|f(y) - f(x)| \le n^{1/q} L \|x - y\|_p .$$

We can use this fact to show that for any $\omega \in [0, 2\pi]$ and any $a', a'' \in A_p$, we have

$$|P_{a'}(e^{j\omega}) - P_{a''}(e^{j\omega})| \le n^{1/q} M_T \|a' - a''\|_p$$

which obviously implies that

$$\|P_{a'} - P_{a''}\|_\infty \le n^{1/q} M_T \|a' - a''\|_p$$

This means that one can choose $\delta^* \le \eta/(n^{1/q} M_T)$ and this completes the proof. □

5 Rank One Mixed μ Synthesis

In this section we briefly review the rank one synthesis (ROS) result of Rantzer and Megretski [23] and specialize it to the fixed–pole model (FPM).

5.1 Fixed–Pole Model as Perturbed Coprime Factors

We first show that the fixed–pole model (FPM) having a hypercube for the parameter uncertainty set is a special case of the perturbed coprime model (PCM). The PCM is the model used in the rank one synthesis theory and is of the form (SISO case)

$$G_{\delta,\Delta} = \frac{N + \delta^T N_\delta + \Delta N_\Delta}{M + \delta^T M_\delta + \Delta M_\Delta}$$

with $N, M \in RH_\infty$, (N, M) coprime, $\delta \in \mathbb{R}^m$, $\|\delta\|_\infty < 1$, $\Delta \in RH_\infty$, $\|\Delta\|_\infty < 1$, $N_\delta, M_\delta \in RH_\infty^m$, $N_\Delta, M_\Delta \in RH_\infty$, $M_\delta = 0$, and $M_\Delta = 0$. The fixed–pole model we want is given by

$$G_{\theta,\Delta} = \frac{B(\theta)}{A} + W\Delta$$

where $B(\theta) = \sum_k \theta_k z^k$, $\theta \in \Theta \subset \mathbb{R}^m$, Θ is a hypercube which is centered at θ_c and has side lengths $\{\eta_k\}$, and $\|\Delta\| \le 1$. This corresponds to the PCM model above with $M = 1$, $N = B(\theta_c)/A$, $N_\Delta = W$, $M_\Delta = 0$, $M_\delta = 0$, and

$$N_\delta = \frac{[\eta_0 \ \eta_1 z \cdots \eta_{m-1} z^{m-1}]^T}{A(z)}$$

We now incorporate the robust performance objective which is given by $\|W_p S\|_\infty \leq \gamma$. This can be transformed into a robust stability problem as follows. Define $G_\gamma \equiv G_{\delta,\Delta}(1 - \Delta_p W_p)^{-1}$ which can also be expressed in terms of the uncertain coprime model as

$$G_a = \frac{N + \delta^T N_\delta + \Delta N_\Delta + \Delta_p N_{\Delta_p}}{M + \delta^T M_\delta + \Delta M_\Delta + \Delta_p M_{\Delta_p}} \tag{8}$$

with $N_{\Delta_p} = 0$, $M_{\Delta_p} = W_p$, and the rest of the quantities defined as above. The goal is to find the largest γ^{-1} such that G_a can be stabilized for all $\|\Delta\| < 1$ and $\|\Delta_p\| < \gamma^{-1}$. Figure 3 shows the model and makes the rank one structure apparent.

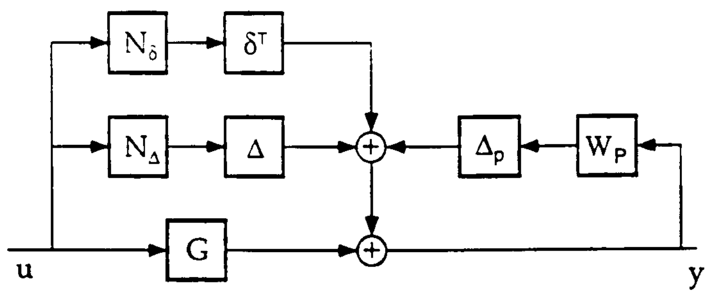


Figure 3: Rank One Structure of the Model

5.2 Solution of the Rank One Synthesis Problem

The general rank one synthesis result was solved by Rantzer and Megretski [23] who derived a convex parameterization of all robustly stabilizing controllers for rank one uncertainty lying in a convex set. We will state this result in the form specialized for the FPM. Let the nominal model, $B(\theta_c)/A$, be denoted by G and define

$$\phi(\alpha,\beta) = \sup_{\omega\in[0,2\pi]} \frac{|W_p([\alpha+\beta]G+\alpha)|_{(e^{i\omega})}}{Re\{\alpha(e^{i\omega})\} - \|Re\{N_\delta[\alpha+\beta]\}_{(e^{i\omega})}\|_d - |N_\Delta(\alpha+\beta)|_{(e^{i\omega})}}$$

where α is a positive real transfer function and β is any stable transfer function. The robust control performance problem described in the preceding section is equivalent to minimizing the functional $\phi(\alpha+\beta)$ over α and β. Although it is a tedious exercise, one can show that the above functional is indeed quasiconvex in (α,β). The denominator being nonnegative is also

a convex constraint and so we get a quasiconvex optimization subject to a convex constraint. An ϵ-optimal solution can be found in finite time using standard methods for convex optimization [2].

6 FPM Step Two: Inner Loop

Having mapped the priors into a FPM set and having a solution to the scheme robust control problem, one can try using an iterative based on the FPM and ROS with a performance objective which is implied by keeping the weighted sensitivity transfer function small. We now describe such a scheme in detail.

The partitioning will be performed with respect to Θ while $W\Delta$ is assumed to represent the inherent nonparametric uncertainty and will remain fixed in size. Thus, we will refer to Θ as the model set and suppress the $W\Delta$ part which is fixed for each parameter value in Θ.
The iterative procedure based on ROS consists of the following steps.

1. Label the initial model set Θ_0 and set $k = 0$.
2. Can the desired performance be achieved for Θ_k by some K_k? If yes, go to (4).
3. Refine Θ_k in the following way (to achieve better performance):
   (a) Find j such that the performance is most sensitive with respect to the j^{th} parameter, θ_j.
   (b) Split Θ_k along the j^{th} dimension, resulting in the two sets X_0 and X_1, with $\Theta_k = X_0 \cup X_1$.
   (c) (Skip if $k = 0$) If X_0 is smaller than the smallest allowable partition size we invalidate Θ_k by decrementing k by 1, and go to (2).
   (d) Find $q \in \{0, 1\}$ such that the best performance which can be achieved for X_q is better than the one for X_{1-q}. Let K_{k+1} be the controller which delivers this performance to X_q.
   (e) Set $\Theta_k = X_{1-q}$, $\Theta_{k+1} = X_q$, increment k by 1, and go to (2).
4. Connect K_k to the plant and test for performance
5. If the performance is satisfied, stop.

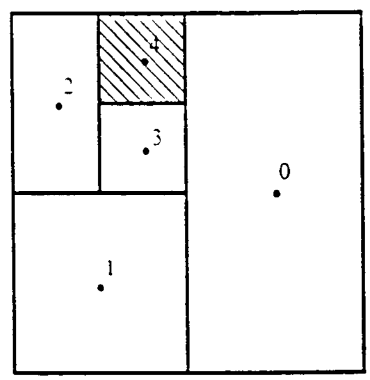


Figure 4: 2D Iteration Example

6. If $k > 0$ invalidate Θ_k by decrementing k by 1 and go to (2). Otherwise, choose a new model parameterization and go to (1).

This procedure has several nice properties. As discussed in Section 3.2, choosing the smallest allowable partition size to be nonzero, we are guaranteed termination in finite time. Every time a set is split, the memory requirement is only increased by one unit (containing the center and side lengths information, for example) so there is no geometric or exponential explosion in required memory. The search is optimistic, always seeking the best set in the partition. At first thought it seems that this may potentially exhibit very bad worst-case behavior. For if the only controller which achieves the performance for the actual process is one that is designed for a "bad" set, the "good" sets will have to be invalidated first. However, the "good" sets will be invalidated quickly because they will typically be larger and will not need to be split as many times as the "bad" sets. The following figure illustrates how the iterations might proceed in the case when Θ has dimension 2. In this example, the shaded box 4 is invalidated, the counter is decremented from 4 to 3, and the procedure resumes by focusing on box 3.

The computationally difficult steps are steps 2, 3d, and possibly 3a. Note that in steps 2 and 3d, we are trying to synthesize controllers meeting either the desired or the best possible performance levels, with step 3d having to solve two such problems. Step 3a which computes the sensitivity of performance with respect to each parameter is fairly easy to compute in the special case of ROS and FPM. The solution is given by the following result.

Lemma 6.1 *Assume that the ROS solution gives a feasible pair $(\alpha(z), \beta(z))$ as well as the worst frequency, ω_0 which maximizes the functional. Then the parameter which has the greatest impact on performance is given by $\theta_{k_{\max}}$, where*

$$k_{max} = \arg \max_{k\in[0,n-1]} \left| Re \left\{ e^{i\omega_0 k} \left(\frac{\beta(e^{i\omega_0}) + \alpha(e^{i\omega_0})}{A(e^{i\omega_0})} \right) \right\} \right| \tag{9}$$

Proof. Given the feasible pair $(\alpha(z), \beta(z))$ and the worst frequency, ω_0, ϕ can be viewed as a function depending only on the η_k's. Thus, we can write

$$\phi(\eta) = \frac{a_0}{a_1 - \|Re\,(N_\delta[\beta + \alpha])\,(e^{i\omega_0})\|_1}$$

where a_0 and a_1 are two real constants. Recall that

$$N_\delta(z) = \frac{[\eta_0 \;\; \eta_1 z \;\; \cdots \;\; \eta_{m-1} z^{m-1}]^T}{A(z)} .$$

Using simple calculus, one can show that

$$\left| \frac{\partial \phi}{\partial \eta_k} \right| \propto \frac{\partial \|Re\,(N_\delta[\beta + \alpha])\,(e^{i\omega_0})\|_1}{\partial \eta_k} .$$

Writing out $\|Re\,(N_\delta[\beta + \alpha])\,(e^{i\omega_0})\|_1$ explicitly and using the fact that $\eta_k > 0$ gives

$$\begin{aligned} \frac{\partial}{\partial \eta_k} \|Re\,(N_\delta[\beta + \alpha])\,(e^{i\omega_0})\|_1 &= \frac{\partial}{\partial \eta_k} \sum_{k=0}^{m-1} \left| Re \left\{ \eta_k e^{i\omega_0 k} \frac{\beta(e^{i\omega_0}) + \alpha(e^{i\omega_0})}{A(e^{i\omega_0})} \right\} \right| \\ &= \left| Re \left(e^{i\omega_0 k} \frac{\beta(e^{i\omega_0}) + \alpha(e^{i\omega_0})}{A(e^{i\omega_0})} \right) \right| \end{aligned}$$

and the result follows. □

6.1 Worst Case Complexity

In this section we consider some issues related to complexity of the iterative scheme and derive some worst–case bounds on the number of experiments and ROS designs. The types of complexity results we are after address the worst case behavior of the scheme with respect to the number of computations and required time for the scheme to terminate. The behavior really

depends on the desired performance level, γ^*, and the initial model set. We begin with the following notation.

Assume that the initial set $\Theta_0 \subset R^m$ is a hypercube with side lengths s_h. Let the *desired* smallest partition size be r_{des}. Next define the number of resolution levels

$$q = \left\lceil \log_2 \frac{s_h}{r_{des}} \right\rceil$$

and so the *actual* smallest partition size is given by $r_p = s_h 2^{-q}$. At each level $k \in [0, q]$ we have a partition Σ_k of Θ_0 consisting of 2^{km} subsets.

We now define $\#D$ as the total number of ROS designs and $\#E$ as the total number of experiments performed. As before, we assume that the existence of a controller means that such a controller can result from a ROS design. We can now state the main results.

Theorem 6.2 *The total number of experiments is bounded as*

$$\#E \leq 2^{mq}$$

with equality only if γ^* cannot be achieved for any partition Σ_k, $k \leq (q-1)$, but can be achieved for **every** *subset in Σ_q.*

Proof. The equality is easy to see. If γ^* can be achieved for some subset $U_j \notin \Sigma_q$, then we perform an experiment for U_j and so $\#E$ is incremented by 1. However, at the end, we either quit, or invalidate U_j, but in either case we will not perform experiments on at least two more sets from Σ_q that are contained in U_j. Thus, we see that at any level $k < q$, we subtract from the worst case $\#E$ if we perform an experiment and therefore the worst case $\#E$ occurs if we perform experiments on all, and only, the subsets of Σ_q. □

Theorem 6.3 *Given the definitions above,*

$$\#D \leq 2^{mq+1} - 1$$

Moreover, there exists a (difficult enough) performance objective, $\gamma^* > 0$ s.t. equality holds.

Proof. Clearly, the worst-case occurs when the performance is so difficult that it cannot be achieved for any subset in Σ_q. This means that eventually, for every subset of every partition, a ROS design will have to be performed.

But these are not the only sets that will have to be operated on. Going from partition Σ_k to Σ_{k+1} requires $m2^{km}$ splits (m for each set), and each split causes the number of designs to increase by two. Since we are considering the case where the entire set will be invalidated, we can reorder the whole procedure and assume that the splits are done in such a way that the resulting sets are in the σ-Algebra generated by Σ_{k+1}. This means that the total number of designs is given by

$$\#D \leq \sum_{k=0}^{qm} 2^k = 2^{mq+1} - 1$$

□

This number is larger than if we had simply considered all the sets in each Σ_k. This is given by $(2^{m(q+1)} - 1)/(2^m - 1)$.

7 Illustrative Example

In this section we present an example which illustrates the iterative scheme described in the previous sections. It is assumed that the plant is known to consist of a second order lightly damped mode with two flexible modes at higher frequencies. The lightly damped mode is known to be of the following form.

$$P(s) = \frac{1}{s^2 + 2\xi\omega_n s + \omega_n^2}$$

where $\underline{\omega} \leq \omega_n \leq \overline{\omega}$, $\underline{\xi} \leq \xi \leq \overline{\xi}$, and it is known that $\underline{\omega} = 0.7$, $\overline{\omega} = 0.8$, $\underline{\xi} = 0.2$, and $\overline{\xi} = 0.3$. It is known that the two other modes occur at frequencies of approximately 8 and 12 rad/s. We can discard these modes and represent them with unmodeled dynamics of the form $W\Delta$. The next two figures show the full and simplified plant. After converting everything to discrete time ($T_{sam} = 0.15$s) we model the simplified plant by the fixed pole approximation. The simplified plant is given by the following.

$$P(a;z) = \frac{0.0121z + 0.0119}{z^2 + a_1 z + a_0} \quad \begin{array}{lcl} a_1 & \in & [-1.9456, -1.9122] \\ a_0 & \in & [0.9274, 0.9570] \end{array}$$

This means that the fixed pole model is of the form

$$G(\theta, \Delta) = \frac{\sum_{k=0}^{m-1} \theta_k z^k + (\epsilon^* + DW)\Delta}{(z^2 - 1.9289z + 0.9422)^{\lfloor m/2 \rfloor}}$$

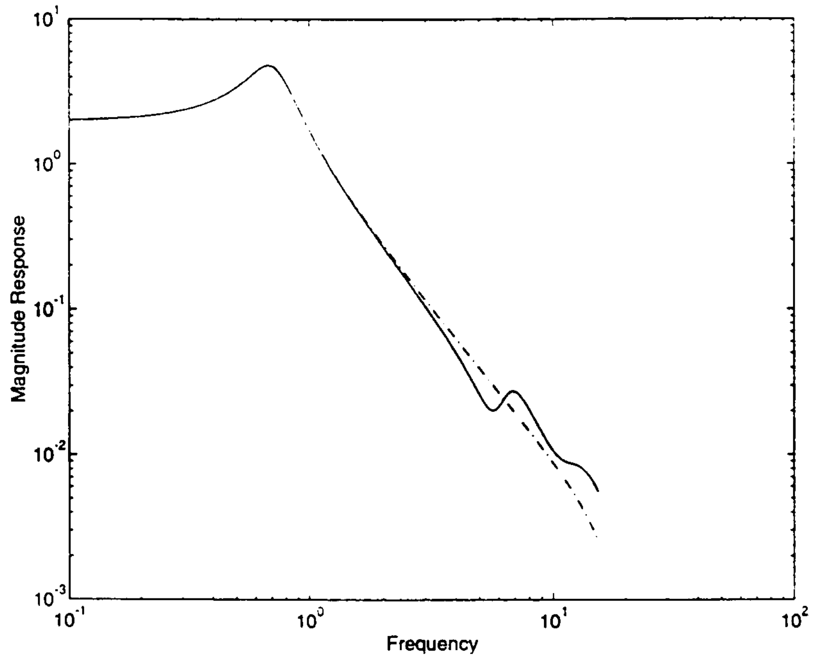


Figure 5: Full Order and Simplified Plant

where ϵ^* is obtained from Section 4.1 and $D(z)$ is the denominator polynomial $(z^2 - 1.9289z + 0.9422)^{\lfloor m/2 \rfloor}$. The ϵ^* represents the error between the FPM and simplified plant, while the DW term is the error between the full order and simplified plant.

We now assume that the actual plant is given by

$$P(z) = \frac{0.00181z^5 - 0.0026z^4 + 0.0055z^3 + 0.0087z^2 + 0.0005z + 0.0053}{z^6 - 2.4015z^5 + 2.5568z^4 - 1.8369z^3 + 0.9620z^2 - 0.4987z + 0.2362}.$$

For visual presentation purposes we consider a fixed pole model of order $m = 2$. The corresponding ϵ^* is 0.0084. The ideal performance objective is assumed to be good tracking of certain duration step inputs, however, for the robust control design we will use small l_2 gain of the weighted sensitivity transfer function as the design objective. We consider a weighting function which will allow designs of fairly demanding bandwidth (i.e., beyond the first lightly damped mode). This weighting function is given by

$$W_p = \frac{z - .2152}{z - .9984}.$$

7.1 Iterative Scheme Simulation Results

We now demonstrate a few examples of the iterative scheme. We take the two dimensional parametric uncertainty set to be the smallest hypercube bounding the set which is asymptotically given by the robust set membership identification algorithm given in [18]. This also serves to show the potential improvement in performance which can be achieved by the iterative scheme. To get an idea of how the scheme might proceed, we compute a number of ROS controllers for various grid refinements and test all of these controllers on the plant. This only gives a rough idea of what the scheme might do because we only consider squares, not rectangles, which will arise due to the sets being split in one dimension at a time. This, however gives a global picture which shows how the predicted robust performance base on the models compares with the actual performance achieved with the plant. These grids are shown in the following figures. The numbers inside the model boxes correspond to the minimum γ achieved for those boxes, while the numbers inside the corresponding plant boxes show the actual γ (i.e., $\|WS\|_\infty$) that those controllers achieve for the actual plant. Note that if one asks for performance level $\gamma > 1.66$, the predicted and actual performance are fairly well correlated and the optimistic search is extremely efficient. However, one can see that as desired performance improves, we must go to

finer grids where the performance is not as well correlated with the predicted robust performance and some invalidation will occur.

Grid 1

3.71	1.66
294	3.11

Actual Performance

1.24	1.16
294	1.26

Grid 2

2.28	1.67	1.40	1.23
3.43	2.15	1.56	1.31
6.92	3.13	1.95	1.43
294	6.19	2.87	1.80

Actual Performance

1.16	1.16	1.24	1.301
1.20	1.13	1.09	1.16
1.27	1.20	1.11	1.05
210	1.37	1.26	1.14

Figure 6: Grids Showing Predicted and Achieved Performance

We now show the evolution of the iterative scheme for various desired performance levels. γ_{des}. This is shown in Figures 8 through 11. The smallest partition size is chosen such that the space is split at most three times (Grid 3 of Figure 7 is the finest allowed partition). The lightly shaded boxes are the ones under current consideration and the dark shaded boxes are the ones that have been invalidated. The values γ correspond to the achievable robust performance for the lightly shaded box, while γ_p corresponds for the performance level achieved when this controller is applied to the actual process. The first execution of the scheme uses $\gamma_{des} = 2.5$ and the results are shown in Figure 8.

The evolution agrees with the data from Grid 1 in Figure 6 and one can

Grid 3

1.95	1.69	1.525	1.425	1.33	1.27	1.21	1.14
2.23	1.87	1.63	1.48	1.38	1.30	1.22	1.17
2.64	2.15	1.79	1.57	1.42	1.335	1.25	1.19
3.28	2.53	2.07	1.73	1.50	1.37	1.30	1.21
4.52	3.13	2.41	1.99	1.66	1.43	1.33	1.25
7.06	4.40	3.01	2.28	1.88	1.58	1.395	1.28
15.68	6.62	4.18	2.85	2.20	1.815	1.53	1.35
294	15.01	6.42	3.99	2.71	2.10	1.71	1.44

Actual Performance

1.17	1.16	1.19	1.25	1.28	1.34	1.38	1.40
1.14	1.13	1.14	1.16	1.22	1.26	1.30	1.34
1.13	1.11	1.08	1.095	1.12	1.185	1.22	1.26
1.16	1.12	1.10	1.06	1.05	1.09	1.16	1.19
1.23	1.17	1.12	1.10	1.055	1.03	1.07	1.115
1.32	1.27	1.18	1.12	1.09	1.05	1.05	1.07
2.16	1.35	1.28	1.20	1.16	1.11	1.06	1.06
210	1.92	1.43	1.34	1.26	1.16	1.11	1.06

Figure 7: Grids Showing Predicted and Achieved Performance

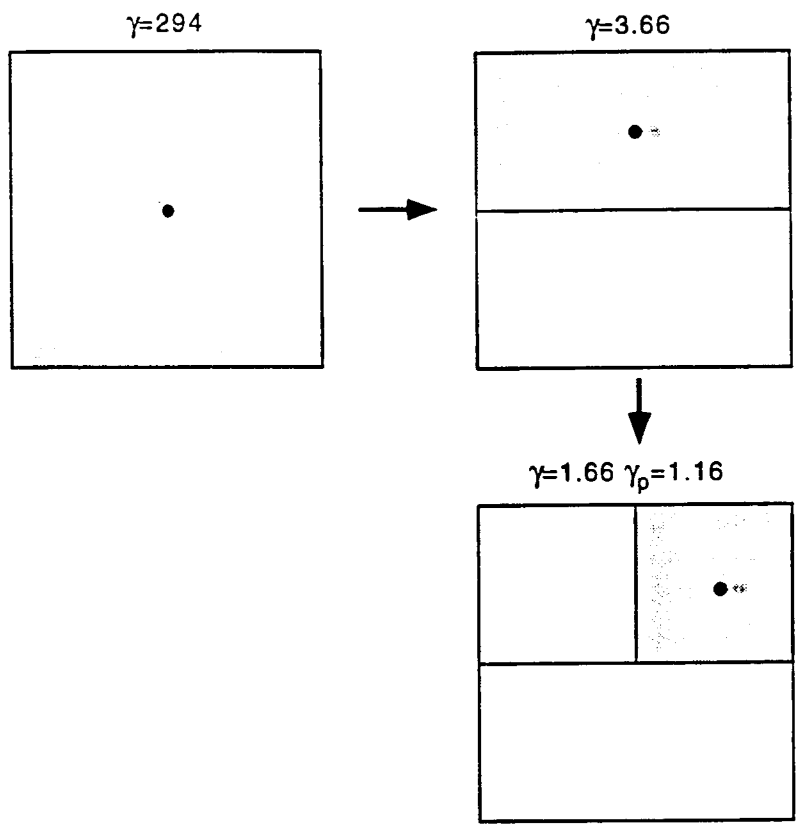


Figure 8: Iterative Scheme ($\gamma_{des} = 2.5$)

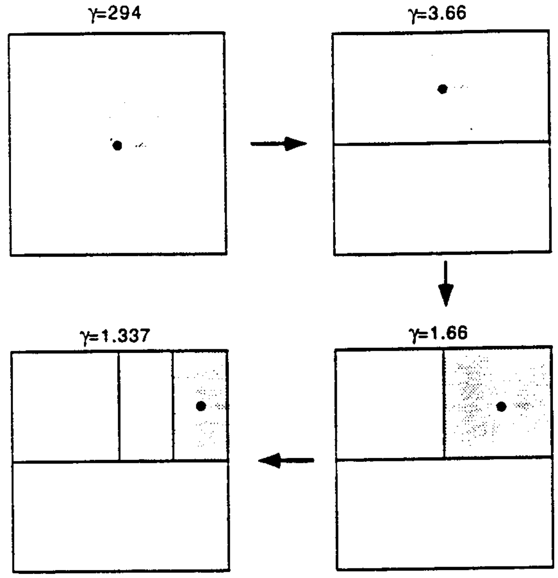


Figure 9: Iterative Scheme: steps 1–4 ($\gamma_{des} = 1.25$)

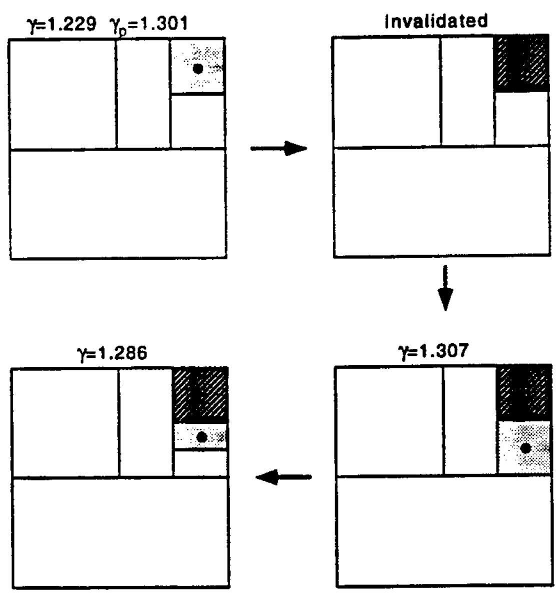


Figure 10: Iterative Scheme: steps 5–8 ($\gamma_{des} = 1.25$)

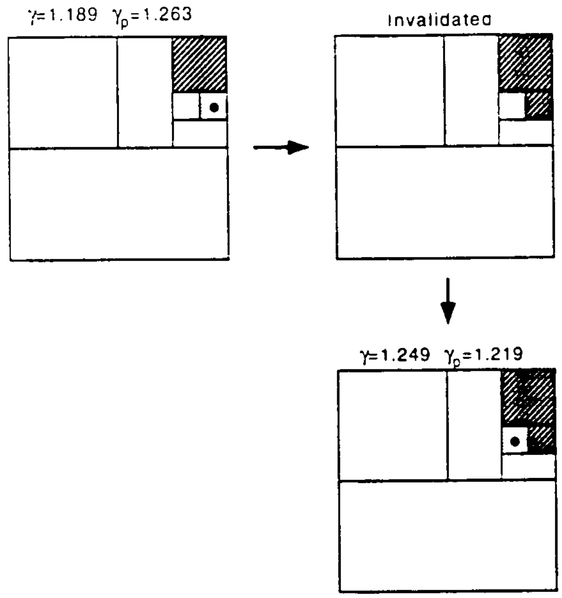


Figure 11: Iterative Scheme: steps 9–12 ($\gamma_{des} = 1.25$)

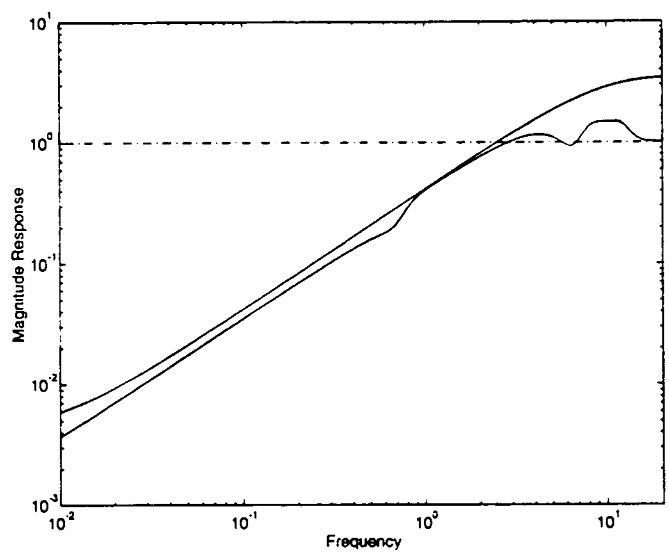


Figure 12: Actual Sensitivity Plot for Final Controller

see that the algorithm marches towards the box which predicts $\gamma = 1.66$ and the actual performance is satisfied ($\gamma_p = 1.16$). The next example considers the case when $\gamma_{des} = 1.25$. In this case, shown in Figures 9 through 11, the optimistic search leads to the upper right corner of Grid 2 and we see that the actual performance will miss the desired value of 1.25 so the upper right corner is invalidated. In addition, the box (two boxes down from upper right box) having $\gamma = 1.189$ is invalidated because the actual performance misses 1.25. Finally, we come to the box which has $\gamma = 1.249$, and the actual performance is met ($\gamma_p = 1.22$). The sensitivity plot of the actual performance and the one guaranteed by the controller for the final box is shown in Figure 12. When the desired performance is $\gamma = 1.2$, the entire model set is invalidated. To achieve better performance than this one can use the 4th order model. Running the iterative scheme with this model and a desired γ of 1.0, the scheme terminates after 11 iterations with a predicted $\gamma = 0.989$ and the achieved $\gamma_p = 0.987$.

8 Discussion

The framework proposed in this paper very general in the sense that it is valid for any mutually consistent model parameterization, robust control design and performance objective. The model set transformation problem for the FPM and a special prior model set (i.e., stable and uncertain poles) was considered. This problem may also be formulated using the gap metric [8] which would allow unstable prior model sets. At this time a scheme based on ROS may be the least conservative because of the lack of conservatism in the rank one synthesis solution. There is, however, a price to be paid for using a simple model such as the FPM. First, the ROS with performance objective given in Section 5.1 is limited to multi-input single-output (MISO) systems. Second, forcing the prior information to be mapped into a FPM can introduce conservatism in the form of large unmodeled dynamics, which will limit achievable peformance. This conservatism can be reduced if one is willing to partition the unmodeled dynamics as well. In other words, one can try to extend the scheme based on ROS and FPM by considering the model $\{\theta^T H + \epsilon\Delta\}$ where ϵ is not fixed but can also be invalidated. One must be careful in this case since the sets $\{\epsilon\Delta : \epsilon_1 \in [0, \eta)\}$ and $\{\epsilon\Delta : \epsilon_1 \in (\eta, \epsilon_1)\}$ are not disjoint even though the values of ϵ are disjoint. Finally, if the entire model set is invalidated one can either change the performance objective or the model. The examples illustrate that increasing the complexity of

the model allows one to achieve better performance through the iterative scheme.

9 Conclusion

This paper presented a new framework for iterative modeling and control. The philosophy of this framework is a very different way of viewing models and their role in designing controllers for uncertain systems. The model is viewed as a tool used to describe the unknown process and really depends on prior information. available control design tools and other modeling preferences. The approach described in this paper was an iterative procedure for refining the uncertainty set via robust control based model invalidation and can be viewed as a systematic way of efficiently searching for a controller delivering a certain desired level of performance to the unknown process. In this way it is possible to invalidate the model if it does not facilitate design of a controller which also provides good performance for the actual process. The result of an iterative scheme in this framework is that either the performance goal will be met or the entire uncertainty set will be invalidated in accordance with our modeling and control method prejudice. An iterative scheme based on a special fixed pole model structure and rank one mixed μ synthesis control design was described in detail and a specific example was used to illustrate the proposed scheme.

10 Acknowledgments

This work was supported in part by the Airforce Office of Scientific Research under Grant AFOSR-91-0368. by the National Science Foundation under Grant NSF 9157306-ECS. and by C.S. Draper Laboratory under Grant DL-H-467128 and Draper IR& D Project No. 438.

References

[1] B. Anderson and R. Kosut. "Adaptive Robust Control: On–Line Learning". pages 297–298. Brighton. England. December 1991.

[2] S.P. Boyd and C.H. Barratt. *Linear Controller Design: Limits of Performance.* Prentice-Hall. 1991.

[3] S.P. Boyd. L. El Ghaoui. E. Feron. and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory.* SIAM. 1994.

[4] M.A. Dahleh and J. Doyle. "From Data to Control". In *Proc. Workshop on Modeling of Uncertainty in Control Systems.* Springer–Verlag, 1992.

[5] M.A. Dahleh and M. Khammash. "Controller Design for Plants with Structured Uncertainty". *Automatica.* 21, January 1993.

[6] J. Doyle and G. Stein. "Beyond Singular Values and Loop Shapes". *Journal of Guidance and Control.* 14(1):5–16, January 1991.

[7] N. Elia. P. Young, and M.A. Dahleh. "Robust Performance for Fixed Inputs". In *Proc. 1994 Conference on Decision and Control*, Orlando. FL.. December 1994.

[8] T. Georgiou and M. Smith. "Optimal Robustness in the Gap Metric". *IEEE Transactions on Automatic Control*, 35:673–686, June 1990.

[9] M. gevers. "Connecting Identification and Robust Control: A New Challenge". Technical Report Rep. 91.48, Université Catholique de Louvain, Louvain, Belgium, 1991.

[10] G. Gu and P. Khargonekar. "Linear and Nonlinear Algorithms for Identification in H^∞ With Error Bounds". *IEEE Transactions on Automatic Control*, 37(7):953–964, July 1992.

[11] A.J. Helmicki. K. Jacobson. and C. Nett. "Control Oriented System Identification in H_∞". *IEEE Transactions on Automatic Control.* 36(10):1163–1176. October 1991.

[12] M. Khammash. "Robust Steady State Tracking". In *Proc.1994 American Control Conf.*, Baltimore, MD., 1994.

[13] P. Khargonekar. "System Identification in Frequency Domain: Theory and Examples". *Proceedings Conf. Feedback Control, Nonlinear Systems, and Complexity*, May 1994.

[14] R. Kosut. M. Lau, and S. Boyd. "Set-Membership Identification of Systems with Parametric and Nonparametric Uncertainty". *IEEE Trans. on Auto. Control*, 37(7):929–941, July 1992.

[15] J. Krause, P.P. Khargonekar, and G. Stein. "Robust Adaptive Control: Stability and Asymptotic Performance". *IEEE Transactions on Automatic Control*, 37(3):316–331, March 1992.

[16] J. Krause, G. Stein, and P.P. Khargonekar. "Robust Performance of Adaptive Controllers with General Uncertainty Structure". In *Proc. 1990 Conference on Decision and Control*, Honolulu, Hawaii, December 1990.

[17] W. Lee, B. Anderson, R. Kosut, and I. Mareels. "On Adaptive Robust Control and Control-Relevent System Identification". In *Proc. 1992 American Control Conference*, pages 2834–2841, Chicago, IL, June 1992.

[18] M. Livstone. *Identification, Robust Adaptation and Iterative Schemes*. PhD thesis, MIT, Cambridge, Massachusetts, October 1994.

[19] J.M. Maciejowski. *Multivariable Feedback Design*. Addison–Wesley, 1989.

[20] P.M. Mäkilä . "Robust Identification and Galois Sequences". Technical Report Rep. 91-1, Åbo Akademi (Swedish University of Åbo), Åbo, Finland, January 1991.

[21] A. nesterov and Y. Nemirovski. *Interior Point Methods in Convex Optimization*. SIAM, 1994.

[22] K. Poolla, P. Khargonekar, A. Tikku, J. Krause, and K. Nagpal. "A Time–Domain Approach to Model Validation". *IEEE Transactions on Automatic Control*, 39(5):351–359, May 1994.

[23] A. Rantzer and A. Megretski. "A Convex Parameterization of Robustly Stabilizing Controllers". Technical report, The Royal Institute of Technology, Stockholm, Sweden, 1993.

[24] R. Schrama and P. Van den Hof. "An Iterative Scheme for Identification and Control Design Based on Coprime Factorizations". In *Proc. 1992 American Control Conference*, pages 2842–2846, Chicago, IL., June 1992.

[25] R.P.J. Schrama. "Accurate Models for Control Design: The Necessity of an Iterative Scheme". *IEEE Transactions on Automatic Control*, 22(2):173-179, July 1992.

[26] R.S. Smith and J.C. Doyle. "Model Invalidation–A Connection Between Robust Control and Identification". *Proc. 1989 American Control Conference*. pages 1435–1440. June 1989.

[27] R.S. Smith and J.C. Doyle. "Model Validation: A Connection Between Robust Control and Identification". *IEEE Transactions on Automatic Control*. 37(7):942–952. July 1992.

[28] D. Tse. M.A. Dahleh. and J. Tsisiklis. "Optimal Asymptotic Identification Under Bounded Disturbances". *Proc. 1992 American Control Conference. Chicago, IL*. pages 679–685. July 1992.

[29] P.M. Young. *"Robustness with Parametric and Dynamic Uncertainty"*. PhD thesis. California Institute of Technology, Pasadena.CA. May 1993.

[30] Z. Zang, R. Bitmead. and M. Gevers. "H_2 Iterative Model Refinement and Control Robustness Enhancement". In *Proc. 1991 Conference on Decision and Control*, pages 279–284, Brighton, England. December 1991.

[31] Z. Zang, R. Bitmead, and M. Gevers. "Disturbance Rejection: On–Line Refinement of Controllers by Closed Loop Modelling". Technical Report Rep. 92.15, Université Catholique de Louvain, Louvain, Belgium, 1992.

[32] K. Zhou. J.C. Doyle. and K. Glover. *H_∞ Control Theory*. Preprint, 1994.

A Framework for Robust Parametric Set Membership Identification

Mitchell M. Livstone* and Munther A. Dahleh†

Abstract

This paper proposes a new framework for studying robust parametric set membership identification. We derive some new results on the fundamental limitations of algorithms in this framework, given a particular model structure. The new idea is to quantify uncertainty only with respect to the (finite dimensional) parametric part of the model and not the (fixed size) unmodeled dynamics. Thus, the measure of uncertainty is different from the measures used in previous robust identification work where system norms are used to quantify uncertainty. As an example, the results are used to assess the fidelity of a certain approximate robust parametric set membership identification algorithm.

1 Introduction

In the past half decade there has been much research activity in the area of robust system identification, otherwise known as control–oriented and control–relevant system identification. The motivation can be attributed to new advances in robust control theory which did not interface well with the existing theory of classical system identification. In particular, robust control requires the plant to be described by a nominal model perturbed by some bounded uncertainty which may or may not have structure. This uncertain *set* of systems is assumed to contain the true plant and the robust control theory provides methods for synthesizing controllers which achieve certain performance, robustly, for the entire uncertainty set. This model set requirement is not satisfied by the classical identification algorithms which typically fix a parametric model structure and then perform some kind of

*Graduate Student, EECS Dept., Massachusetts Institute of Technology
†Associate Professor, EECS Dept., Massachusetts Institute of Technology

regression to get a value for the parameters. This yields a *single*, finite dimensional, identified system. Thus, the main focus of current research in robust identification has been the formulation of algorithms that yield nominal plants along with measures of uncertainty which are well suited for existing robust control methodologies; hence, the terms control-oriented and control-relevant. Since these algorithms yield uncertain *sets* of plants, all of these robust identification algorithms can be classified as some kind of set membership identification (SMID) algorithms.

The formulations and algorithms essentially differ in the types of *a priori* information assumed about the model set and disturbances. The model set assumption is partially driven by the description of modeling uncertainty required by the robust control design. The frequency domain algorithms in [7, 6] provide nominal models along with unstructured uncertainty which is bounded in H_∞ and thus provides the correct description for H_∞ control theory [5, 3]. The time domain algorithms in [11, 13] provide nominal models with uncertainty bounds in the l_1 norm and are well suited for the l_1 control theory [2]. Of course, the l_1 norm provides a (potentially conservative) bound for the H_∞ norm so the time domain algorithms can also be used in H_∞ robust control, however conservative the bounds may be.

The algorithms mentioned above are formulated in a worst-case asymptotic setting. The formulation is worst-case with respect to the plant and noise. In other words, given the worst allowable noise and the worst plant in the original model set, the identified set must contain the true plant. Furthermore, the algorithm should be asymptotically convergent in the sense that in the limit as the noise $\rightarrow 0$, cardinality of data and the nominal plant order both $\rightarrow \infty$, the worst-case identification error (i.e., distance between nominal and true plants) goes to zero.

More recently, some robust extensions of the parametric set membership identification setup have appeared in the literature [15, 9, 8]. The roots of parametric set membership identification (PSMID) can be traced back to the late 1960's in the work of Schweppe [12] and Bertsekas [1] who studied state estimation under unknown but bounded disturbances. These ideas were later applied to system identification (parameter estimation) by Fogel [4] and a steady flow of papers on SMID has persisted ever since. The models used in these papers were simple ARMA models with some output additive (unknown but bounded) noise. Since these models are of fixed finite dimension, they are not very useful for robust control. This motivated Younce, Krause and others [15, 9, 8] to consider a model with unstructured uncertainty. Algorithms which use this model set will be referred to as ro-

bust parametric SMID (RPSMID) and an example of a prior model set is the following set with additive unstructured uncertainty.

$$\mathcal{M}_0 = \{G_\theta + W\Delta : \theta \in \Theta_0 \subset \mathbb{R}^m, \|\Delta\|_\infty \leq 1\} \tag{1}$$

where $G_\theta(z)$ is a SISO, rational transfer function whose polynomial coefficients are elements of the parameter vector θ, and W is a known (assumed) weighting function for the uncertainty.

Surprisingly, essentially all of the work in PSMID over the last 15 years has focused on the construction of algorithms and computations, with very little mention of convergence issues. A formal framework which can address issues such as fundamental limitations, uncertainty and optimal inputs seems to be missing. In the special case of FIR models, the work of Tse [13] can be applied, although it yields conservative results.

In this paper we introduce a new, robust parametric set membership identification (RPSMID) framework and derive some results on the fundamental limitations of algorithms in this framework. The development is similar to the work of Tse, et.al. [13], and the results can be viewed as generalizations of some of the results therein. This paper is organized as follows. The RPSMID problem is formulated in the following section. The reset of the paper is concerned with the diameter of the uncertainty set and optimal inputs which can shrink the uncertainty set to its theoretical minimum. In Sections 3–5 we present some results on the size of the worst–case uncertainty sets and optimal inputs for two special cases: noise–free ($\delta = 0$) and purely parametric ($\Delta = 0$). Section 6 contains the corresponding results for the general case. Section 7 illustrates how these results can be used to assess the fidelity of an approximate RPSMID algorithm.

2 Problem Formulation

Let the linear time invariant plant model set be given by:

$$\mathcal{M}_0 = \{G(\theta, \Delta) : \theta \in \Theta_0 \subset \mathbb{R}^m, \|\Delta\| \leq 1\} \tag{2}$$

where Δ is defined on some Banach space (H_∞ or l_1). Given an input $u \in l_p$ and $\|u\|_p \leq 1$, the experiment is defined by

$$y = h_p * u + d \qquad h_p \in \mathcal{A}, \|d\| \leq \delta \tag{3}$$

We assume that the parameters to be identified are given by θ, while the inherent unstructured uncertainty in the model is captured by Δ. The exact

model structure is embedded in the functional form of G. For example, an additive uncertainty structure can be expressed by

$$G(\theta, \Delta) = \frac{B_\theta}{A_\theta} + W\Delta \ .$$

A set identification algorithm, ϕ, maps the experiment data up to time n and the priors to an identified set $\mathcal{M}_n$:

$$\phi(P_n(u, y), \Delta, \delta, \Theta_0) = \mathcal{M}_n$$

The plant membership set, $\mathcal{S}_n$ is given by all plants consistent with the observed data (up to time n) and the priors.

$$\begin{aligned} \mathcal{S}_n \quad = \quad & \{G(\theta, \Delta) : P_n(y - G(\theta, \Delta) * u - d) = 0 \\ & \text{for some } \theta \in \Theta_0, \|\Delta\| \leq 1, \|d\| \leq \delta\} \end{aligned} \tag{4}$$

The parametric membership set, Θ_n^c, is the set of all parameter values consistent with the data and priors, and is given by

$$\Theta_n^c = \{\theta \in \mathbb{R}^m : G(\theta, \Delta) \in \mathcal{S}_n\} \tag{5}$$

Note that $\mathcal{S}_n \subseteq \{G(\theta, \Delta) : \theta \in \Theta_n^c, \|\Delta\| \leq 1\} \equiv \tilde{\mathcal{S}}_n$, which is actually the uncertainty set of plants generated by Θ_n^c and all possible Δ in the unit ball.

An algorithm is said to be RPSMID when it satisfies

$$\mathcal{S}_n \subseteq \mathcal{M}_n \ .$$

This simply means that the identified set contains the consistent set. The fidelity of the algorithm can then be judged based on how tight an overbound it provides for the consistent set.

In the RPSMID framework, the identification error for any RPSMID algorithm is bounded from below by the diameter of $\tilde{\mathcal{S}}_n$. However, since the size of Δ is fixed, it only makes sense to study the convergence of the size of the parametric set Θ_n^c.

We now define uncertainty in this framework and examine its asymptotic worst-case behavior. Uncertainty will be measured *only with respect to the parametric part* of the model. Thus, the uncertainty associated with the set $\mathcal{S}_n$ will be the same as the uncertainty in $\tilde{\mathcal{S}}_n$. This uncertainty is defined as the diameter of Θ_n^c with respect to the metric ρ on $\mathbb{R}^m$ defined in the following way.

$$\text{diam}(\Theta) = \sup_{\theta_1 \in \Theta} \sup_{\theta_2 \in \Theta} \rho(\theta_1, \theta_2) \tag{6}$$

It will be understood that (with slight abuse of notation) $\text{diam}(\mathcal{S}_\infty)$ means $\text{diam}(\Theta^c_\infty)$ as defined above. Using this definition we define the worst–case diameter of uncertainty in the standard way.

$$D(u, \mathcal{M}_0, \Delta, \delta) = \sup_{h \in \mathcal{M}_0} \sup_{\|d\| \leq \delta} \text{diam}(\mathcal{S}_\infty(\mathcal{M}_0, u, u * h + d, \Delta, \delta))$$

At this level of generality it is difficult to say anything about the diameter of uncertainty. We will impose more structure and try to derive more specialized results. In particular, we will consider the additive uncertainty structure

$$G(\theta, \Delta) = G(\theta) + W\Delta$$

where $G(\theta)$ is affine in θ and $\Theta_0 = \theta_c + \tilde{\Theta}_0$, where $\tilde{\Theta}_0$ is a balanced and convex set in $\mathbb{R}^m$. Furthermore, we require that W is stable and has a stable inverse in the space where it will be defined. These assumptions will hold for the remainder of this chapter.

3 Purely Parametric Case

We first consider the case where $\Delta = 0$ so the uncertainty enters only through the disturbance, d. We derive lower bounds for the diameter of uncertainty and a few special cases. Later we consider the existence of inputs which can asymptotically shrink the uncertainty to these lower bounds. The main result is given by the following theorem which provides a lower bound for the diameter of the uncertainty set.

Theorem 3.1 *If $\exists \theta_1, \theta_2 \in \Theta_0$ (closed and bounded) s.t. $\|G(\theta_1) - G(\theta_2)\|_1 \geq 2\delta$, then for any input $u \in l_p$,*

$$D(u, \mathcal{M}_0, 0, \delta) \geq 2 \sup \{\|\theta - \theta_c\| : \|G(\theta) - G(\theta_c)\|_1 \leq \delta\}$$

where θ_c is the analytic center of Θ_0

$$\inf_{\theta_1 \in \mathbb{R}^m} \sup_{\theta \in \Theta_0} \|\theta - \theta_1\| = \sup_{\theta \in \Theta_0} \|\theta - \theta_c\|$$

Remark 3.2 *The norms in the theorem are not specified because any norm defined on $\mathbb{R}^m$ can be used to measure θ and the diameter is defined with respect to the metric induced by this norm. The norm used to measure the operators is the same as the norm used for Δ.*

Before proving the theorem we need a lemma and a fact. We can assume, without loss of generality, that $G(\theta)$ is linear (not affine) in θ.

Lemma 3.3

$$\sup_{h\in\mathcal{M}_0}\sup_{\|d\|\le\delta} diam(\Theta^c_\infty(\mathcal{M}_0, u, u*h+d, \delta)) = diam(\Theta^c_\infty(\mathcal{M}_0, u, u*G(\theta_c), \delta))$$

Proof. The above means that given any $\theta_p \in \Theta_0$, $\|d\| \le \delta$ and $\epsilon > 0$, **if** $\exists\theta_1, \theta_2 \in \Theta^c_\infty(\mathcal{M}_0, u, u*G(\theta_p)+d, \delta)$ and $\rho(\theta_1,\theta_2) = \epsilon$, then $\exists\theta_1', \theta_2' \in \Theta^c_\infty(\mathcal{M}_0, u, u*G(\theta_c), \delta)$ and $\rho(\theta_1', \theta_2') \ge \epsilon$. We will now show that one can always choose $\theta_1' = \theta_c + \frac{\theta_1-\theta_2}{2}$ and $\theta_2' = \theta_c - \frac{\theta_1-\theta_2}{2}$. We first show that $\theta_1', \theta_2' \in \Theta_0$ by showing that $\frac{\tilde\theta_1-\tilde\theta_2}{2} \in \tilde\Theta_0$. Because $\tilde\Theta_0$ is a balanced set, θ_1 and $-\theta_2$ are both in it. Combining it with the fact that $\tilde\Theta_0$ is also convex gives the result $\theta_1', \theta_2' \in \Theta_0$. We now show that θ_1' and θ_2' are also in $\Theta^c_\infty(\mathcal{M}_0, u, u*G(\theta_c), \delta)$. Define $\tilde\theta = \theta_1 - \theta_2$. Since $\theta_1, \theta_2 \in \Theta^c_\infty(\mathcal{M}_0, u, u*h+d, \delta)$ implies

$$h*u+d = G(\theta_1)*u+d_1 = G(\theta_2)*u+d_2$$

for some $\|d_1\| \le \delta$ and $\|d_2\| \le \delta$, this means that

$$G(\tilde\theta)*u = d_2 - d_1 .$$

Now, if the plant is given by $h = G(\theta_c)$ and $d = 0$ we have

$$y = G(\theta_c)*u = G(\theta_c + \frac{\tilde\theta}{2}) + *u + d_1' = G(\theta_c - \frac{\tilde\theta}{2}) + *u + d_2'$$

which implies that $G(\frac{\tilde\theta}{2})*u = -d_1'$ and $G(\frac{\tilde\theta}{2})*u = d_2'$. This shows that one can take $d_1' = \frac{1}{2}(d_1-d_2)$ and $d_2' = -d_1'$. Since $\|d_1'\| = \|d_2'\| \le \delta$, this shows that $\theta_1', \theta_2' \in \Theta^c_\infty(\mathcal{M}_0, u, u*G(\theta_c), \delta)$. Furthermore, $\rho(\theta_1', \theta_2') = \rho(-\frac{\tilde\theta}{2}, \frac{\tilde\theta}{2}) = \epsilon$.
□

Fact 3.4 *If Θ is a balanced and convex set, then*

$$\sup_{\theta_1,\theta_2\in\Theta}\|\theta_1-\theta_2\| = 2\sup_{\theta\in\Theta}\|\theta\|$$

Proof. (Thm) Lemma 3.3 above shows that the worst case uncertainty occurs when the plant is $G(\theta_c)$ and $d = 0$. Then given any $\|u\| \le 1$, for

each $\theta \in \Theta_\infty^c(\mathcal{M}_0, u, u * G(\theta_c), \delta)$ we have $\|(G(\theta_c) - G(\theta)) * u\| \leq \delta$ which is implied by $\|G(\theta_c) - G(\theta)\|_1 \leq \delta$. This means that

$$\{\theta : \|G(\theta_c) - G(\theta)\|_1 \leq \delta\} \subseteq \Theta_\infty^c(\mathcal{M}_0, u, u * G(\theta_c), \delta)$$

for any $\|u\| \leq 1$. This implies that for any $\|u\| \leq 1$

$$D(u, \mathcal{M}_0, 0, \delta) = \operatorname{diam}(\Theta_\infty^c(\mathcal{M}_0, u, u * G(\theta_c), \delta)) \geq \operatorname{diam}\left(\hat{\Theta}^c\right)$$

and since $\{\theta : \|G(\theta_c) - G(\theta)\|_1 \leq \delta\}$ is a (θ_c) translation of a balanced and convex set, $\operatorname{diam}\left(\hat{\Theta}^c\right) = 2 \cdot \sup\left\{\|\theta - \theta_c\| : \theta \in \hat{\Theta}^c\right\}$ (by Fact 3.4). □

Corollary 3.5 *If* $G(\theta) = \frac{\sum(\theta_k)z^{-k}}{A(z)}$ *and diam is measured w.r.t. the* 1*-norm, then*

$$D(u, \mathcal{M}_0, 0, \delta) \geq 2\delta/\beta$$

where $\beta = \|1/A\|_1$

Note that this bound may not be tight, but as $A \to 1$ it becomes tight and agrees with a result by Tse, et.al. [14]. It is also interesting to compare this approximation to the one which would be obtained by applying Tse's result directly to this case. Tse's result can be applied here when G is FIR so it is necessary to multiply through by $A(z)$. After defining $\tilde{y} = Ay$, the experiment becomes

$$\tilde{y} = g * u + A * d \quad \|d\|_\infty \leq \delta \ .$$

Now one can get an inner approximation to the uncertainty set:

$$\{g : \|g * u - \tilde{y}\|_\infty \leq \|A\|_1\delta\} \subseteq \{g : \exists \|d\|_\infty \leq \delta \text{ s.t. } \|g * u - \tilde{y}\|_\infty \leq \|Ad\|_\infty\}$$

and so $D(u, \mathcal{M}_0, 0, \delta) \geq 2\|A\|_1\delta$. Comparing this result to the corollary, one can see that this result is at least as conservative as the approximation in the corollary. This is due to the fact that given any invertible operator, A, and an induced norm $\|\cdot\|$, the inequality

$$\|A\| \geq \frac{1}{\|A^{-1}\|}$$

always holds.

4 Noise Free Case

We now consider the case where $\delta = 0$ so the inherent uncertainty enters only through Δ. The main result gives a lower bound for the diameter of the uncertainty set.

Theorem 4.1 *If* $\exists \theta_1, \theta_2 \in \Theta_0$ *s.t.* $\left\| \frac{G(\theta_1)-G(\theta_2)}{W} \right\| \geq 1$, *then for any input* $u \in l_p$

$$D(u, \mathcal{M}_0, \Delta, 0) \geq 2 \sup \left\{ \|\theta - \theta_c\| : \left\| \frac{G(\theta) - G(\theta_c)}{W} \right\| \leq 1 \right\} \tag{7}$$

where θ_c *is the analytic center of* Θ_0.

Before proving the theorem we need one lemma. We again assume, without loss of generality, that $G(\theta)$ is linear (not affine) in θ.

Lemma 4.2

$$\sup_{h \in \mathcal{M}_0} diam(\Theta_\infty^c(\mathcal{M}_0, u, u * h, \Delta)) = diam(\Theta_\infty^c(\mathcal{M}_0, u, u * G(\theta_c), \Delta))$$

Proof. The proof is similar to the proof of Lemma 3.3 and is therefore omitted. □

Proof. (Thm) Lemma 4.2 above shows that the worst case plant is given by $\Delta = 0$ and $\theta_{\text{act}} = \theta_c$. Once this is shown, assume that the plant is $G(\theta_c)$. Then given any input, if $G(\theta_c) = G(\theta_1) + W\Delta$ for some $\|\Delta\| \leq 1$, it follows that $\theta_1 \in \Theta_\infty^c$. This shows that

$$\left\{ \theta : \left\| \frac{G(\theta) - G(\theta_c)}{W} \right\| \leq 1 \right\} \subseteq \Theta_\infty^c$$

but since this set is convex and balanced (with translation), Fact 3.4 shows that

$$\text{diam} \left(\left\{ \theta : \left\| \frac{G(\theta) - G(\theta_c)}{W} \right\| \leq 1 \right\} \right) = 2 \sup \left\{ \|\theta - \theta_c\| : \left\| \frac{G(\theta) - G(\theta_c)}{W} \right\| \leq 1 \right\}$$

This completes the proof. □

We now consider a special case.

Corollary 4.3 *If $G = \frac{\sum \theta_k z^{-k}}{A(z)}$, the diameter and Δ are measured w.r.t. the l_1 norm on Ξ^m and l_1, respectively.*

$$D(u, \mathcal{M}_0, \Delta, 0) \geq \frac{2}{\beta}$$

where $\beta = \|1/AW\|_1$.

Proof. We simply note that

$$\left\| \frac{\sum(\theta_k)z^{-k}}{A(z)W(z)} \right\|_1 \leq \left\| \frac{1}{AW} \right\|_1 \| \sum_{k=0}^{m-1} (\theta_k) z^{-k} \|_1 = \left\| \frac{1}{AW} \right\|_1 \|\theta\|_1$$

This means that

$$\left\{ \theta : \left\| \frac{1}{AW} \right\|_1 \|\theta\|_1 \leq 1 \right\} \subseteq \left\{ \theta : \left\| \frac{\sum(\theta_k)z^{-k}}{A(z)W(z)} \right\|_1 \leq 1 \right\}$$

and the result follows after applying Theorem 4.1. □

5 Optimal Inputs

In this section we show that there exist inputs which can decrease the diameter of the uncertainty set to the theoretical lower bounds derived in the previous sections. We use Galois sequences and arguments similar to those of Tse, et.al [14] and Mäkilä [11]. A Galois sequence of order n is a minimum length binary sequence which contains every possible subsequence of length $\leq n$. We consider the parametric case first.

5.1 Optimal Inputs – Parametric Case

In the purely parametric case ($W = 0$) the theoretical lower bound is given by

$$2 \sup \{\|\theta - \theta_c\| : \|G(\theta) - G(\theta_c)\|_1 \leq \delta\}$$

We can assume WLOG that $\theta_c = 0$ and define

$$\alpha = 2 \sup \{\|\theta\| : \|G(\theta)\|_1 \leq \delta\}$$

The main result shows that this lower bound is in fact tight when a Galois sequence input, u^* is used.

Theorem 5.1 *Let all definitions of Section 3 hold. Then*

$$D(u^*, \mathcal{M}_0, 0, \delta) \leq \alpha$$

Proof. The theorem will be proved for G linear (not affine) in θ but the extension to the affine case is trivial. Choose any $\theta \in \Theta_\infty^c(\mathcal{M}, u^*, 0, \delta)$. This implies that $\|G(\theta) * u^*\|_\infty \leq \delta$, and since $G(\theta)$ is linear in θ

$$G(\theta) = \sum_{k=0}^{N} g_k(\theta) z^{-k} \quad g_k(\theta) \text{ linear in } \theta \ \forall k$$

Since the input u^* is a Galois sequence, $\exists m > 0$ s.t.

$$[u_m^* \ u_{m+1}^* \ldots u_{m+N}^*] = [\mathrm{sgn}(g_N(\theta)) \ \mathrm{sgn}(g_{N-1}(\theta)) \ldots \mathrm{sgn}(g_0(\theta))]$$

This means that

$$\begin{aligned} \left|(G(\theta) * u^*)_{m+N}\right| &= \left|\sum_{k=m}^{m+N} g_{k-m}(\theta) u_{m+N-k}\right| \\ &= \sum_{k=m}^{m+N} |g_{k-m}(\theta)| = \|G(\theta)\|_1 \end{aligned}$$

Putting this together gives

$$\delta \geq \|G(\theta) * u^*\|_\infty \geq \left|(G(\theta) * u^*)_{m+N}\right| = \|G(\theta)\|_1$$

This means that $\|\theta\| \leq \alpha/2$ and thus $D(u^*, \mathcal{M}_0, 0, \delta) \leq \alpha$ □

5.2 Optimal Inputs – Noise Free Case

Recall that in this case the experiment is given by

$$y = (G(\theta) + W\Delta) * u$$

and assume without loss in generality that the set Θ_0 is centered at zero ($\theta_c = 0$). Define the theoretical lower bound in Theorem 4.1 (with $\theta_c = 0$) as

$$\alpha = 2 \sup \left\{ \|\theta\| : \left\| \frac{G(\theta)}{W} \right\| \leq 1 \right\}$$

The main result shows that this lower bound is in fact tight when a Galois sequence input, u^* is used.

Theorem 5.2 *Let all definitions of Section 4 hold. Then*

$$D(u^*, \mathcal{M}_0, \Delta, 0) \leq \alpha$$

Proof. The proof considers the linear case, but it is easy to see that the result holds for affine G as well. Choose any $\theta \in \Theta^c_\infty(\mathcal{M}, u^*, \Delta, 0)$. This implies that

$$\left\| \frac{G(\theta)}{W} * u^* \right\|_\infty \leq \|\Delta * u^*\|_\infty$$

Furthermore, define $\tilde{G}(\theta) = \frac{G(\theta)}{W}$ which belongs to l_1 and is also linear in θ. This gives

$$\tilde{G}(\theta) = \sum_{k=0}^{\infty} \tilde{g}_k(\theta) z^{-k} \quad \tilde{g}_k(\theta) \text{ linear in } \theta \ \forall k$$

and for any $\epsilon > 0$, $\exists N > 0$ s.t. $\sum_{k=N+1}^{\infty} |\tilde{g}_k| < \epsilon/2$. Since the input u^* is a Galois sequence, $\exists m > 0$ s.t.

$$[u^*_m \ u^*_{m+1} \dots u^*_{m+N}] = [\text{sgn}(\tilde{g}_N(\theta)) \ \text{sgn}(\tilde{g}_{N-1}(\theta)) \dots \text{sgn}(\tilde{g}_0(\theta))]$$

This means that for every $\epsilon > 0$, $\exists m, M > 0$ s.t.

$$\begin{aligned}
\left| \left(\tilde{G}(\theta) * u^*\right)_{m+N} \right| &= \left| \sum_{k=0}^{m+N} \tilde{g}_k(\theta) u_{m+N-k} \right| \\
&= \left| \sum_{k=0}^{N} \tilde{g}_k(\theta) u_{m+N-k} + \sum_{k=N+1}^{N+m} \tilde{g}_k(\theta) u_{m+N-k} \right| \\
&= \left| \sum_{k=0}^{N} \text{sgn}(\tilde{g}_k(\theta)) \tilde{g}_k(\theta) + \sum_{k=N+1}^{N+m} \tilde{g}_k(\theta) u_{m+N-k} \right| \\
&\geq \sum_{k=0}^{N} |\tilde{g}_k(\theta)| - \sum_{k=N+1}^{N+m} |\tilde{g}_k(\theta)| \\
&\geq \|\tilde{G}(\theta)\|_1 - \epsilon
\end{aligned}$$

Putting this together gives

$$\begin{aligned}
\left\| \frac{G(\theta)}{W} \right\| &\leq \left| \left(\frac{G(\theta)}{W} * u^* \right)_{m+N} \right| \leq \left\| \frac{G(\theta)}{W} * u^* \right\|_\infty \\
&\leq \|\Delta * u^*\| \leq \|\Delta\|_1 \|u^*\|_\infty \leq \|\Delta\|_1 \leq 1
\end{aligned}$$

This means that $\|\theta\| \leq \alpha/2$ and thus $D(u^*, \mathcal{M}_0, 0, \delta) \leq \alpha$ □

6 General Model

We are now ready to consider the general case with uncertainty as well as a disturbance. This case is more difficult and we cannot get an exact expression which is not a function of Δ or d. One can again show that the worst case situation is $\theta = \theta_c$, $\Delta = 0$, and $d = 0$. We will assume WLOG that $\theta_c = 0$. This gives the following expression.

$$D(u, \mathcal{M}_0, \Delta, \delta) \geq 2 \sup \{\|\theta\| : \|G(\theta) + W\Delta\| \leq \delta \text{ for some } \|\Delta\| \leq 1\}$$

We can use an argument entirely similar to the one used in the two previous Sections on optimal inputs, and show that using a Galois input we can match the sign of any N consecutive elements of the impulse response of $G(\theta) + W\Delta$. This will show that for this optimal input, the set

$$\Theta^*(\Delta, \delta) \equiv \{\theta : \|G(\theta) - W\Delta\| \leq \delta \text{ for some } \|\Delta\| \leq 1\}$$

is equivalent to the consistent parameter set Θ^c_∞. We can similarly define the corresponding sets for the two special cases

$$\Theta^*(\Delta, 0) \equiv \left\{\theta : \left\|\frac{G(\theta)}{W}\right\| \leq 1\right\}$$

and

$$\Theta^*(0, \delta) \equiv \{\theta : \|G(\theta)\|_1 \leq \delta\}$$

We can now show the following result.

Lemma 6.1

$$\Theta^*(\Delta, 0) \oplus \Theta^*(0, \delta) \subseteq \Theta^*(\Delta, \delta) \tag{8}$$

where $\oplus$ is the Minkowski set addition.

Proof. Choose any $\theta_1 \in \Theta^*(\Delta, 0)$ and $\theta_2 \in \Theta^*(0, \delta)$. This means that for any $\|u\| \leq 1$, there is some $\|\Delta\| \leq 1$ and some $\|d\| \leq \delta$ such that $[G(\theta_1) + W\Delta] * u = 0$ and $G(\theta_2) * u - d = 0$. But this implies that $[G(\theta_1 + \theta_2) + W\Delta] * u - d = 0$ which means that $\theta_1 + \theta_2 \in \Theta^*(\Delta, \delta)$. □

Unfortunately, the diameter of this set sum cannot be tightly bounded from below. In fact, it is easy to derive the following bounds.

$$\max(\mathrm{diam}(\Theta^*(\Delta,0)),\mathrm{diam}(\Theta^*(0,\delta))) \leq \mathrm{diam}(\Theta^*(\Delta,0)\oplus\Theta^*(0,\delta))$$

and

$$\mathrm{diam}(\Theta^*(\Delta,0)\oplus\Theta^*(0,\delta)) \leq \mathrm{diam}(\Theta^*(\Delta,0)) + \mathrm{diam}(\Theta^*(0,\delta))$$

We can also show that

$$\Theta^*(\Delta,\delta) \subseteq \{\theta : \|G(\theta)\|_1 \leq \delta + \|W\|_1\} \tag{9}$$

This gives the following bounds for $\mathrm{diam}(\Theta^*(\Delta,\delta))$.

Theorem 6.2

$$\max(diam(\Theta^*(\Delta,0)), diam(\Theta^*(0,\delta))) \leq diam(\Theta^*(\Delta,\delta))$$

and

$$diam(\Theta^*(\Delta,\delta)) \leq diam(\{\theta : \|G(\theta)\|_1 \leq \delta + \|W\|_1\})$$

These bounds are not tight except when $W \to 0$. In fact, when $\delta \ll \|W\|_1$ the upper bound can be very poor since the set $\{\theta : \|G(\theta)\|_1 \leq \|W\|_1\}$ can be much larger than $\left\{\theta : \left\|\frac{G(\theta)}{W}\right\|_1 \leq 1\right\}$.

7 Assessing Fidelity of Approximate Algorithms

In this section we show an application of the results developed in the previous sections. In particular we use these results to study the conservatism of a particular approximate ellipsoid RPSMID algorithm presented in [10]. The analysis in [10] presents worst-case asymptotic uncertainty results under optimal inputs. The algorithm itself is not important and we only state the relevant result. Given the model structure

$$G(\theta,\Delta) = G(\theta) + W\Delta$$

where $G(\theta)$ is affine in θ and the experiment is given by

$$y = G(\theta,\Delta)u + d \quad, \|d\|_\infty \leq \delta$$

the asymptotic worst-case diameter of the approximate parametric membership set (an ellipse) measured in the ∞ norm on $\mathbb{R}^m$ is given by

$$\text{diam(Ellipse)} \leq 2\|A\|_1(\|W\|_1 + \delta)$$

We will use this in conjunction with the exact asymptotic results developed in this paper as follows. The worst-case exact and approximate results can be stated roughly as

$$\text{diam(exact)} \geq v \quad \text{and} \quad \text{diam(approx)} \leq w$$

and this implies that diam(approx)$\leq \frac{w}{v}$ diam(exact). This gives an upper bound for the diameter of the approximate solution as a function of the exact diameter of uncertainty. The main result is as follows.

Theorem 7.1 *For the model given by*

$$y = \frac{B(\theta)}{A}u + \Delta W u + d$$

the worst-case diameter of uncertainty under Galois inputs is given by

$$\mathit{diam(ellipse)} \leq (1 + \|W\|_1\|W^{-1}\|_1)(\|A\|_1\|A^{-1}\|_1)\,\mathit{diam}\,(\Theta^*(\Delta, \delta))$$

Proof. From [10] we know that the diameter (under optimal inputs) measured in the ∞-norm satisfies

$$\text{diam(Ellipse)} \leq 2\|A\|_1(\|W\|_1 + \delta) \tag{10}$$

We now derive an explicit lower bound for the diameter of the exact uncertainty set using Lemma 6.1. Notice that the proof of Lemma 6.1 is valid whether Δ is measured in l_1 or H_∞. We make the distinction between these two norms more explicit by defining the two sets

$$\Theta^*_{(\infty)}(\Delta, 0) \equiv \left\{\theta : \left\|\frac{G(\theta)}{W}\right\|_\infty \leq 1\right\}$$

and

$$\Theta^*_{(1)}(\Delta, 0) \equiv \left\{\theta : \left\|\frac{G(\theta)}{W}\right\|_1 \leq 1\right\}$$

while the definition

$$\Theta^*(0, \delta) \equiv \{\theta : \|G(\theta)\|_1 \leq \delta\}$$

holds as in Lemma 6.1. It is easy to see that $\Theta^*_{(1)}(\Delta, 0) \subseteq \Theta^*_{(\infty)}(\Delta, 0)$. The above observation and definitions can easily be used to show that when $\Theta^*(\Delta, \delta)$ is defined with $\|\Delta\|_\infty \leq 1$ we have

$$\Theta^*_{(1)}(\Delta, 0) \oplus \Theta^*(0, \delta) \subseteq \Theta^*_{(\infty)}(\Delta, 0) \oplus \Theta^*(0, \delta) \subseteq \Theta^*(\Delta, \delta) \qquad (11)$$

We also have

$$\operatorname{diam}\left(\Theta^*_{(\infty)}(\Delta, 0)\right) \geq \frac{2}{\|(AW)^{-1}\|_\infty} \geq \frac{2}{\|(AW)^{-1}\|_1} \geq \frac{2}{\|A^{-1}\|_1 \|W^{-1}\|_1}$$

and

$$\operatorname{diam}\left(\Theta^*(0, \delta)\right) \geq \frac{2\delta}{\|A^{-1}\|_1}$$

which we can combine to get

$$\operatorname{diam}\left(\Theta^*(\Delta, \delta)\right) \geq \frac{2}{\|A^{-1}\|_1} \max\left(\frac{1}{\|W^{-1}\|_1}, \delta\right) . \qquad (12)$$

We now combine Equations 12 and 10 to get

$$\operatorname{diam}(\text{Ellipse}) \leq \frac{\delta + \|W\|_1}{\max\left(\frac{1}{\|W^{-1}\|_1}, \delta\right)} \|A\|_1 \|A^{-1}\|_1 \operatorname{diam}\left(\Theta^*(\Delta, \delta)\right)$$

The result follows from the fact that

$$\begin{aligned} \frac{\delta + \|W\|_1}{\max\left(\frac{1}{\|W^{-1}\|_1}, \delta\right)} &= \frac{\delta}{\max\left(\frac{1}{\|W^{-1}\|_1}, \delta\right)} + \frac{\|W\|_1}{\max\left(\frac{1}{\|W^{-1}\|_1}, \delta\right)} \\ &\leq 1 + 1 \cdot \|W\|_1 \|W^{-1}\|_1 \end{aligned}$$

□

There will typically not be a problem with the W term since it can usually be picked as a fairly well conditioned system. The A terms cannot be controlled as easily. If the poles are close to the unit circle, the quantity $\|A\|_1 \|A^{-1}\|_1$ may be large and the approximate solution may be somewhat conservative in this case.

8 Conclusion

In this paper we have presented a new framework for studying robust parametric set membership identification. Some concrete results concerning the diameter of the worst-case uncertainty set were derived for an affine in the parameters model structure. It was also shown that Galois inputs are optimal for asymptotically shrinking the worst-case diameter of uncertainty. These results were then applied to the assessment of fidelity of a certain approximate robust parametric set membership identification algorithm. It is not known whether similar results can be developed for more sophisticated model structures and this is a direction for future research.

References

[1] D.P. Bertsekas and I.B. Rhodes. "Recursive State Estimation for a Set-Membership Description of Uncertainty System Inputs". *IEEE Transactions on Automatic Control*, 16:117–128, 1971.

[2] M.A. Dahleh and M. Khammash. "Controller Design for Plants with Structured Uncertainty". *Automatica*, 21, January 1993.

[3] J. Doyle, K. Glover, P. Khargonekar, and B. Francis. "State-Space Solutions to Standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Control Problems". *IEEE Transactions on Automatic Control*, 34(8):831–847, August 1989.

[4] E. Fogel. "System Identification via Membership Set Constraints with Energy Bounded Noise". *IEEE Transactions on Automatic Control*, 24(5):752–758, October 1979.

[5] B. Francis. *A Course in H_∞ Control Theory*. Lecture Notes in Control and Information Systems, Vol. 88. Spring-Verlag, 1987.

[6] G. Gu and P. Khargonekar. "Linear and Nonlinear Algorithms for Identification in H^∞ With Error Bounds". *IEEE Transactions on Automatic Control*, 37(7):953–964, July 1992.

[7] A.J. Helmicki, K. Jacobson, and C. Nett. "Control Oriented System Identification in H_∞". *IEEE Transactions on Automatic Control*, 36(10):1163–1176, October 1991.

[8] R. Kosut, M. Lau, and S. Boyd. "Set-Membership Identification of Systems with Parametric and Nonparametric Uncertainty". *IEEE Trans. on Auto. Control*, 37(7):929–941, July 1992.

[9] J. Krause and P.P. Khargonekar. "Parameter Identification in the Presence of Nonparametric Dynamic Uncertainty". *Automatica*, 26(1):113–123, 1990.

[10] M. Livstone. *Identification, Robust Adaptation and Iterative Schemes.* PhD thesis, MIT, Cambridge, Massachusetts, September 1994.

[11] P.M. Mäkilä . "Robust Identification and Galois Sequences". Technical Report Rep. 91-1, Åbo Akademi (Swedish University of Åbo), Åbo, Finland, January 1991.

[12] F.C. Schweppe. "Recursive State Estimation: Unknown but Bounded Errors and System Inputs". *IEEE Transactions on Automatic Control*, 13:22–28, 1968.

[13] D. Tse, M.A. Dahleh, and J. Tsisiklis. "Optimal Asymptotic Identification Under Bounded Disturbances". *Proc. 1992 American Control Conference, Chicago, IL*, pages 679–685, July 1992.

[14] D.N.C. Tse, M.A. Dahleh, and J.N. Tsisiklis. "Optimal Asymptotic Identification Under Bounded Disturbances". *IEEE Transactions on Automatic Control*, 38(8):1176–1190, August 1993.

[15] R. Younce and C. Rohrs. "Identification with Non-Parametric Uncertainty". *Proc. 1990 Conference on Decision and Control, Honolulu, Hawaii*, pages 3154–3161, December 1990.

Set Membership Identification Algorithms and Asymptotic Properties

Mitchell M. Livstone* and Munther A. Dahleh†

Abstract

This paper addresses the asymptotic worst-case properties of set membership identification (SMID) algorithms. We first present a set membership identification algorithm which can be used with a model structure consisting of parametric and nonparametric uncertainty, as well as output additive disturbances that are deterministic and magnitude bounded. This algorithm is then studied in the context of asymptotic worst-case behavior. We derive a lower bound on the worst-case achievable identification error, which is measured by the volume of the identified ellipsoidal uncertainty sets. We then show that there exist inputs which can shrink the uncertainty sets to this lower bound.

1 Introduction

The roots of set membership identification can be traced back to the late 1960's in the work of Schweppe [15] and Bertsekas [2] who studied state estimation under unknown but bounded disturbances. These ideas were later applied to identification (parameter estimation) by Fogel [6] and a steady flow of papers on SMID has persisted ever since [7, 13, 14, 1, 16, 9, 5]. Most of the work in SMID over the last 15 years has focused on the construction of algorithms and computations, with very little mention of convergence issues, especially optimal input design. The work on SMID algorithms can be subdivided mainly into two categories: ellipsoids and polytopes. This work aims at constructing algorithms which tightly bound the parametric uncertainty set with ellipsoids and polytopes, respectively. Given the disturbance assumptions $\|d\|_\infty \leq \delta$, ellipsoid algorithms are more

*Graduate Student, EECS Dept., Massachusetts Institute of Technology
†Associate Professor, EECS Dept., Massachusetts Institute of Technology

conservative than their polytope counterparts since the polytopes give the exact characterization of uncertainty in this case. However, what is lost in conservatism is gained in computational simplicity. The recursive ellipsoid algorithms require only matrix multiplication, while the polytope algorithms typically require solutions to linear programs.

There are essentially two types of model structures considered in the SMID literature. The first is a purely parametric (ARMA) model with additive, unknown but bounded disturbances [6, 7, 13, 14, 1, 5]. This includes a large portion of the SMID work in the literature. The more recent approach uses a model structure which is parametric along with nonparametric uncertainty (either additive or multiplicative), but no disturbance [16, 10]. One exception is the work of Kosut, et.al. [9] which discusses nonparametric uncertainty and disturbances which are either stochastic or satisfy a spectral energy bound. The presence of nonparametric uncertainty does not allow polytope–type algorithms to be used and one is left with either ellipsoid algorithms for approximate characterizations or infinite dimensional convex programs for exact characterizations of uncertainty. In this paper we develop an algorithm for parametric and nonparametric uncertainty with additive, magnitude bounded noise and study its asymprotic properties in detail.

The paper is organized as follows. First we derive a recursive ellipsoid algorithm which can be used for the general model described above. This is very similar to recursive least squares and is a simple extension of some of the ellipsoid algorithms in the literature. In Section 3.2 we study the worst–case asymptotic properties of this algorithm. A lower bound on the worst–case volume is derived for a model structure whose parametric part is linear in the parameters. We next show that using a random binary sequence input will shrink the worst–case volume to this lower bound, with probability one, and using a Galois sequence input will surely shrink the volume asymptotically to this value.

2 Background

Various models have been used in the formulation of the SMID problem. The model used in Fogel [6] is

$$y_k = \theta^T \phi_k + d_k \quad \text{with } \|d\|_\infty \leq \delta \tag{1}$$

where $\theta^T = [\theta_1 \ \theta_2 \cdots \theta_m]$ is the vector of unknown parameters and $\phi = [y_{k-1} \cdots y_{k-p} \ u_k \cdots u_{k-q}]$ is the usual regressor vector.

The set of parameters consistent with the *single* observation at time k is defined by

$$S_k \equiv \left\{ \theta \in \mathbb{R}^m : |y_k - \theta^T \phi_k| \leq \delta \right\} \tag{2}$$

and the set consistent with all observations up to, and including time k, is defined as

$$\Theta_k^* = \bigcap_{l=1}^{k} S_l \tag{3}$$

The *identification goal* is to find a set, Θ_k, at each time k, which satisfies $\Theta_k^* \subseteq \Theta_k$ and the inclusion is as tight as possible, in some appropriate sense. It is interesting to note that only the set Θ_k is updated so the identified and the original model sets have identical structure. Furthermore, the size of the uncertainty set is a function of the experiment. This means that disturbances which are not "worst–case" actually affect the size of parametric uncertainty. This can be seen from the simple example where only one parameter has to be identified (i.e., $y_k = \theta u_k + d_k$). In this case, the parametric uncertainty set is an interval in $\mathbb{R}$. The worst–case disturbance is obviously zero and one can see that a disturbance sequence which takes on values $-\delta$ and $+\delta$ will shrink the uncertainty set to a single point.

Clearly, the exact solution at time k is an intersection of k sets, and each set is defined by two supporting hyperplanes. Much of the research has focused on these exact algorithms [13, 14]. On the other hand, the pointwise inequality in Equation 2 implies

$$\Theta_n^* \subseteq \left\{ \theta : \sum_{k=1}^{n} \lambda_k (y_k - \theta^T \phi_K)^2 \leq \delta^2 \sum_{k=1}^{n} \lambda_k \right\} \tag{4}$$

where $\lambda_k \geq 0$ are free parameters to be picked judiciously. This set defines an ellipsoid and much of the work over the last 15 years has focused on efficient computation and reduction in conservatism of the above bound [6, 7, 4]. In particular, Fogel and Huang [7] derive equations for choosing the parameters λ_k to achieve minimum volume (determinant) or minimum trace ellipsoids. This model structure is not very useful for robust control because it assumes perfect knowledge of the plant order and relative degree. This motivated Younce, Krause and others [16, 10, 9] to consider a model with unstructured uncertainty. The following is a model set with additive unstructured uncertainty.

$$\mathcal{M}_0 = \{G_\theta + W\Delta : \theta \in \Theta_0 \subset \mathbb{R}^m, \|\Delta\|_\infty \leq 1\} \tag{5}$$

where $G_\theta(z)$ is a SISO, rational transfer function whose polynomial coefficients are elements of the parameter vector θ, and W is a known (assumed) weighting function for the uncertainty. The process set is simply given by

$$y = h * u \quad \text{and } h \in \mathcal{M}_0$$

In this case there is no exact characterization of the parametric uncertainty set in terms of polytopes since the pointwise bound $|(\Delta u)_k| \leq |u_k|$ does not necessarily hold, and a more complicated exact characterization can be derived [12]. The approximate characterization can be expressed in terms of ellipsoids.

3 Approximate Set Membership ID

In this section we set up the *approximate* set membership identification problem for a particular model structure. We state the recursive equations for the noise–free case and then extend these to the noisy case.

Let the process model be given by:

$$y = G_{\theta,\Delta} u + \frac{F}{A} d \text{ where} \tag{6}$$

$$G_{\theta,\Delta} \equiv \frac{B_\theta + \Delta W}{A}, \quad \|d\|_{l_\infty} \leq \delta, \ \|\Delta\|_{H_\infty} \leq 1 \ , \tag{7}$$

$$\begin{aligned} A(z) &= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{m-1} z^{-m+1} \\ B_\theta(z) &= \theta_0 + \theta_1 z^{-1} + \theta_2 z^{-2} + \cdots + \theta_{m-1} z^{-m+1} \\ W(z) &= w_0 + w_1 z^{-1} + w_2 z^{-2} + \cdots + w_{m-1} z^{-m+1} \\ F(z) &= f_0 + f_1 z^{-1} + f_2 z^{-2} + \cdots + f_{m-1} z^{-m+1} \end{aligned}$$

We also define $f \equiv [f_0 \ f_1 \cdots f_n]^T$ and $\theta \equiv [\theta_0 \ \theta_1 \cdots \theta_{m-1}]^T$. We next show that the consistent parameter set can be efficiently bounded by an ellipsoid. Let $\tilde{y}_k \equiv y_k + a_1 y_{k-1} + \cdots + a_{m-1} y_{k-m+1} = (Ay)_k$,
$\tilde{u}_k \equiv w_0 u_k + w_1 u_{k-1} + \cdots + w_{m-1} u_{k-m+1} = (Wu)_k$ and
$\phi_k \equiv [u_k \ u_{k-1} \cdots u_{k-m+1}]^T$. Assuming that the data was generated by such a model, at each time k we have

$$\tilde{y}_k - \theta^T \phi_k = (\Delta W u)_k + (Fd)_k \tag{8}$$

In this case, weighting the significance of the data at each time step by

$$\lambda_k^{1/2}(\tilde{y}_k - \theta^T \phi_k) = \lambda_k^{1/2}(\Delta\tilde{u}_k + (Fd)_k), \quad \lambda_k \geq 0 \tag{9}$$

as is done in Fogel and Huang [7] is not possible since the inequality

$$\sum_{k=1}^{N}(\Delta\tilde{u})_k^2 \leq \sum_{k=1}^{N}\tilde{u}_k^2$$

does not hold pointwise. By writing down the running sums we get

$$\begin{aligned} \left\{\sum_{k=1}^{N}(\tilde{y}_k - \theta^T\phi_k)^2\right\}^{1/2} &= \left\{\sum_{k=1}^{N}[(\Delta\tilde{u})_k + (Fd)_k)]^2\right\}^{1/2} \\ &\leq \left\{\sum_{k=1}^{N}\tilde{u}_k^2\right\}^{1/2} + \left\{\sum_{k=1}^{N}(Fd)_k^2\right\}^{1/2} \end{aligned}$$

Noting that $(Fd)_k^2 = (f_0 d_k + f_1 d_{k-1} + \cdots + f_n d_{k-n})^2 \leq \delta^2\|f\|_1^2$ we get the following.

$$\sum_{k=1}^{N}(\tilde{y}_k - \theta^T\phi_k)^2 \leq \Psi_N^2 \quad \text{where} \tag{10}$$

$$\Psi_N^2 \equiv \sum_{k=1}^{N}\tilde{u}_k^2 + \delta^2\|f\|_1^2 N + 2\delta\|f\|_1 N^{1/2}\left\{\sum_{k=1}^{N}\tilde{u}_k^2\right\}^{1/2}$$

3.1 Recursive Equations

In this section we derive recursive equations for the ellipsoid matrix as well as the nominal parameter estimate. We begin by setting $\delta = 0$ and thus $\Psi_N^2 = \sum_{k=1}^{N}\tilde{u}_k^2$. Equation 10 now simplifies to

$$\sum_{k=1}^{N}(\tilde{y}_k - \theta^T\phi_k)^2 \leq \sum_{k=1}^{N}\tilde{u}_k^2$$

For the sake of the derivation we assume that $\sum_{k=1}^{N}\phi_k\phi_k^T$ is invertible, however, this assumption is not necessary for the recursive algorithm to work if appropriate initializations are made. Expanding, and defining $\Gamma_N^{-1} \equiv \sum_{k=1}^{N}\phi_k\phi_k^T$, $\beta_N \equiv \sum_{k=1}^{N}\tilde{y}_k\phi_k$ and $\alpha_N \equiv \sum_{k=1}^{N}(\tilde{y}_k^2 - \tilde{u}_k^2)$, we get the following.

$$\theta^T\Gamma_N^{-1}\theta - 2\theta^T\beta_N + \alpha_N \leq 0$$

Defining $\hat{\theta}_N = \Gamma_N \beta_N$, $V_N = \beta_N^T \Gamma_N \beta_N - \alpha_N$ and $P_N = \Gamma_N V_N$ we get the form for an ellipsoid. The ellipsoidal parameter set at time N is then given by

$$\Theta_N = \left\{ \theta \in R^m \mid (\theta - \hat{\theta}_N)^T P_N^{-1} (\theta - \hat{\theta}_N) \leq 1 \right\} \tag{11}$$

We next derive recursions for $\hat{\theta}$, Γ and V, which requires some algebraic manipulation and the matrix inversion lemma (MIL):

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B + C^{-1})^{-1}DA^{-1}$$

The recursion for Γ_N is just a simple application of MIL with $A = \Gamma^{-1}$, $B = \phi$, $C = 1$ and $D = \phi^T$.

$$\begin{aligned} \Gamma_{N+1}^{-1} &= \Gamma_N^{-1} + \phi_{N+1}\phi_{N+1}^T \\ \Gamma_{N+1} &= \Gamma_N - \frac{\Gamma_N \phi_{N+1} \phi_{N+1}^T \Gamma_N}{1 + \phi_{N+1}^T \Gamma_N \phi_{N+1}} \end{aligned}$$

The equation for $\hat{\theta}_{N+1}$ is simply

$$\hat{\theta}_{N+1} = \Gamma_{N+1}\beta_{N+1} = \Gamma_{N+1}(\beta_N + \tilde{y}_{N+1}\phi_{N+1})$$

The recursion is obtained using a few more steps.

$$\begin{aligned} \hat{\theta}_{N+1} &= \Gamma_{N+1}(\beta_N + \tilde{y}_{N+1}\phi_{N+1}) \\ &= \Gamma_{N+1}(\Gamma_N^{-1}\hat{\theta}_N + \tilde{y}_{N+1}\phi_{N+1}) \\ &= \Gamma_{N+1}([\Gamma_{N+1}^{-1} - \phi_{N+1}\phi_{N+1}^T]\hat{\theta}_N + \tilde{y}_{N+1}\phi_{N+1}) \\ &= \hat{\theta}_N + \Gamma_{N+1}(\tilde{y}_{N+1} - \phi_{N+1}^T\hat{\theta}_N)\phi_{N+1} \end{aligned}$$

The recursion for V_N requires a bit more work and turns out to be

$$V_{N+1} = V_N + \tilde{u}_{N+1}^2 - \frac{(\tilde{y}_{N+1} - \hat{\theta}_N^T \phi_{N+1})^2}{1 + \phi_{N+1}^T \Gamma_N \phi_{N+1}} .$$

The computations for the noise-free case are then as follows.

$$\Gamma_{N+1} = \Gamma_N - \frac{\Gamma_N \phi_{N+1} \phi_{N+1}^T \Gamma_N}{1 + \phi_{N+1}^T \Gamma_N \phi_{N+1}} \tag{12}$$

$$\hat{\theta}_{N+1} = \hat{\theta}_N + \Gamma_{N+1}(\tilde{y}_{N+1} - \phi_{N+1}^T\hat{\theta}_N)\phi_{N+1} \tag{13}$$

$$V_{N+1} = V_N + \tilde{u}_{N+1}^2 - \frac{(\tilde{y}_{N+1} - \hat{\theta}_N^T \phi_{N+1})^2}{1 + \phi_{N+1}^T \Gamma_N \phi_{N+1}} \tag{14}$$

The equations for the noisy case can now be derived as a simple extension of the above.

In the general case we still get an ellipsoid given by

$$\theta^T \Gamma_N^{-1} \theta - 2\theta^T \beta_N + \alpha_N \leq 0$$

with Γ and β unchanged, however, α is now given by

$$\begin{aligned}
\alpha_N &\equiv \sum_{k=1}^{N} \tilde{y}_k^2 - \Psi_N^2 \\
&= \sum_{k=1}^{N} (\tilde{y}_k^2 - \tilde{u}_k^2) - \delta^2 \|f\|_1^2 N - 2\delta \|f\|_1 N^{1/2} \left\{ \sum_{k=1}^{N} \tilde{u}_k^2 \right\}^{1/2}
\end{aligned}$$

Noticing that the recursions for Γ and $\hat{\theta}_N$ do not depend on α, the recursions for the noise–free case hold here as well. The recursion for V_N can be obtained easily from the noise–free case by simply noting that

$$\begin{aligned}
V_{N+1} - V_N &= (\beta_{N+1}^T \Gamma_{N+1} \beta_{N+1} - \beta_N^T \Gamma_N \beta_N - \tilde{y}_{N+1}^2) + \tilde{u}_{N+1}^2 + \delta^2 \|f\|_1^2 \\
&+ 2\delta \|f\|_1 \left\{ \sqrt{N+1} \left(\sum_{k=1}^{N+1} \tilde{u}_k^2 \right)^{1/2} - \sqrt{N} \left(\sum_{k=1}^{N} \tilde{u}_k^2 \right)^{1/2} \right\}
\end{aligned}$$

The part in the parentheses is equivalent to the noise–free case and we end up with the following. Setting $\kappa_0 = 0$, at time $N+1$ we have

$$\kappa_{N+1} = \kappa_N + \tilde{u}_{N+1}^2 \tag{15}$$

$$\begin{aligned}
V_{N+1} &= V_N - \frac{(\tilde{y}_{N+1} - \hat{\theta}_N^T \phi_{N+1})^2}{1 + \phi_{N+1}^T \Gamma_N \phi_{N+1}} + \tilde{u}_{N+1}^2 + \\
&\quad \delta^2 \|f\|_1^2 + 2\delta \|f\|_1 \left\{ \sqrt{N+1}\, \kappa_{N+1}^{1/2} - \sqrt{N}\, \kappa_N^{1/2} \right\}
\end{aligned} \tag{16}$$

3.2 Worst Case Behavior of Ellipsoidal Algorithms

In this section we consider the worst case asymptotic behavior of the ellipsoid algorithm derived in the previous section. We first derive a lower bound on the worst case achievable volume and then show that by using a random binary sequence as an input, this lower bound can be achieved asymptotically in time.

Let the process be modeled by

$$\tilde{y}_k = \theta^T \phi_k + (\epsilon \Delta u)_k + (Fd)_k \;, \quad \|d\|_\infty \leq \delta, \; \|\Delta\|_\infty \leq 1 \tag{17}$$

The next result shows that given any input $\|u\|_\infty \leq 1$, there is a fundamental worst-case lower bound on the volume of the identified ellipsoids. We derive a lower bound on the determinant of the ellipsoid matrix, which is proportional to the square of the volume of the ellipsoid. The exact relationship is given by

$$\mathrm{Vol}(E_N) = \frac{\pi^{m/2}\sqrt{\det(P_N)}}{\Gamma(m/2+1)} \tag{18}$$

where Γ is the Gamma function and for any integer n, $\Gamma(n+1) = n!$ and $\Gamma(1/2) = \sqrt{\pi}$.

Theorem 3.1 *Assuming that the true plant, g, is in $\mathcal{M}$, and the ellipsoids are computed according to Equations* 12 *through* 16,

$$\inf_{\|u\|_\infty \leq 1} \sup_{g\in\mathcal{M}} \sup_{\|d\|_\infty \leq \delta} \det(P_N) \geq (\epsilon + \delta\|f\|_1)^{2m}$$

Proof. Assuming that the original ellipsoid is centered at zero (no loss in generality) and is large enough, one can easily show that the worst case situation occurs when $\theta_{plant} = 0$, $\Delta = 0$, and $d = 0$, which implies that $\tilde{y} = 0$. Subsequent ellipsoids are then given by

$$\Theta_N = \left\{\theta \in R^n \ : \ \theta^T P_N^{-1}\theta \leq 1\right\} \tag{19}$$

where $P_N = V_N\Gamma_N$, $\Gamma_N^{-1} = \sum_k \phi_k\phi_k^T$ and

$$V_N = N\delta^2\|f\|_1^2 + \epsilon^2\sum_{k=0}^{N} u_k^2 + 2\epsilon\delta\|f\|_1\sqrt{N}\left(\sum_{k=0}^{N} u_k^2\right)^{1/2}$$

We will actually show that

$$[\det(P_N)]^{\frac{1}{m}} = V_N\,[\det(\Gamma_N)]^{\frac{1}{m}} \geq (\epsilon + \delta\|f\|_1)^2.$$

Since $\Gamma_N^{-1} = \sum_k \phi_k\phi_k^T$ and $\phi_k = [u_k \cdots u_{k-m+1}]^T$, it is true that

$$\left(\Gamma_N^{-1}\right)_{j,j} = \sum_{k=1}^{N-j+1} u_k^2 \ \text{ for all } 1 \leq j \leq m\ .$$

It is also true that Γ_N^{-1} is positive semidefinite since for any $x \in R^m$

$$x^T\left(\sum_{k=1}^{N}\phi_k\phi_k^T\right)x = \sum_{k=1}^{N} x^T\phi_k\phi_k^T x = \sum_{k=1}^{N} c_k^2 \geq 0$$

We can now apply Hadamard's Inequality [8] to Γ^{-1} to show that

$$\det\left(\Gamma_N^{-1}\right) \leq \prod_{j=1}^{m}\left(\Gamma_N^{-1}\right)_{j,j} = \prod_{j=1}^{m}\left(\sum_{k=1}^{N-j+1} u_k^2\right) \leq \left(\sum_{k=1}^{N} u_k^2\right)^m$$

Using the above facts we now have the following sequence of inequalities.

$$\begin{aligned}
\inf_{\|u\|_\infty\leq 1} \left[\det(P_N)\right]^{\frac{1}{m}} &= \inf_{\|u\|_\infty\leq 1} \frac{N\delta^2\|f\|_1^2 + 2\sqrt{N}\delta\epsilon\|f\|_1\left(\sum_{k=1}^N u_k^2\right)^{1/2} - \epsilon^2\sum_{k=1}^N u_k^2}{\left[\det\left(\sum_k^N \phi_k\phi_k^T\right)\right]^{\frac{1}{m}}} \\
&\geq \inf_{\|u\|_\infty\leq 1} \frac{N\delta^2\|f\|_1^2 + 2\sqrt{N}\delta\epsilon\|f\|_1\left(\sum_{k=1}^N u_k^2\right)^{1/2} + \epsilon^2\sum_{k=1}^N u_k^2}{\sum_{k=1}^N u_k^2} \\
&\geq \inf_{\|u\|_\infty\leq 1} \frac{N\delta^2\|f\|_1^2}{\sum_{k=1}^N u_k^2} + \inf_{\|u\|_\infty\leq 1} \frac{2\sqrt{N}\delta\epsilon\|f\|_1}{\left(\sum_{k=1}^N u_k^2\right)^{1/2}} + \epsilon^2 \\
&\geq (\epsilon + \delta\|f\|_1)^2
\end{aligned}$$

□

We now show that this volume can be achieved if a special input is used. Let $\{u_k\}$ be a random binary sequence obtained by a series of independent Bernoulli trials where

$$\text{Prob}\{u_k = -1\} = \text{Prob}\{u_k = +1\} = 1/2$$

The main result is captured in the following theorem.

Theorem 3.2 *If u is random binary sequence, then*

$$\lim_{N\to\infty} \sup_{g\in\mathcal{M}_0} \sup_{\|d\|_\infty\leq\delta} \det(P_N) = (\epsilon + \delta\|f\|_1)^{2m} \quad i.p.$$

Proof. We now consider Γ more carefully. In particular, recalling that $\phi_k = [u_k\, u_{k-1} \cdots u_{k-m}]^T$, we explicitly write Γ in terms of the input sequence as

$$\left(\Gamma_N^{-1}\right)_{i,j} = \sum_{k=1}^{N} u_{k-i+1}u_{k-j+1}$$

It is apparent that for all $i \leq m$

$$\left(\Gamma_N^{-1}\right)_{i,i} = \sum_{k=1}^{N} u_{k-i+1}^2 = N - i$$

since $u_k = 0$ for $k \leq 0$. Now we show that Γ^{-1} becomes diagonally dominant.

First, notice that taking expectations we get

$$E\left\{\sum_{k=1}^{N} u_{k-i+1}u_{k-j+1}\right\} = \sum_{k=1}^{N} E\left\{u_{k-i+1}u_{k-j+1}\right\}$$
$$= \sum_{k=\max(i,j)}^{N} \delta_{i-j} = (N - \max(i,j))\delta_{i-j}$$

which means that the *expectation* of Γ^{-1} is diagonal. We must now show that the sequence of random variables $\{u_{k-i+1}u_{k-j+1}\}_k$ are independent. To see this, we first note that for a r.v. u_iu_j, fixing u_i does not change the pdf and we get

$$p(u_iu_j|u_j = 1) = p(u_iu_j|u_j = -1) = p_u(u_iu_j)$$

where p_u is the pdf for each of the u_k's. Assuming that $i < j$ (symmetry) and defining $q = j - i$, we rewrite the above sum in the simpler form

$$\left(\Gamma_N^{-1}\right)_{j-q,j} = \sum_{k=1}^{N-j+1} u_{k+q}u_k$$

Now it is clear that for $k < q+1$ the elements are all independent. When $k = q+1$, we are summing the two r.v.'s $u_{2q+1}u_{q+1}$ and $u_{q+1}u_1$. But since u_{2q+1} and u_1 are independent, one can see that $u_{2q+1}u_{q+1}$ and $u_{q+1}u_1$ are also independent. We can use the same argument throughout the sum and show that $\{u_{k+q}u_k\}_k$ is a sequence of independent Bernoulli random variables taking on values in $\{-1,+1\}$, each with probability $1/2$. Each has a mean of zero and $\sigma = 1$. The variance for the sum is then given by

$$\sigma = \sqrt{N-j+1} \leq \sqrt{N}$$

Now we can use Chebyshev's Inequality to show that for $i \neq j$, any $\alpha > 0$ and $\epsilon > 0$

$$\Pr\left\{\frac{\left|\left(\Gamma_N^{-1}\right)_{i,j}\right|}{N^{(\frac{1}{2}+\alpha)}} > \epsilon\right\} \leq \frac{N}{\epsilon^2 N^{(1+2\alpha)}} = \frac{1}{\epsilon^2 N^{2\alpha}}$$

which shows that for each $i \neq j$, the sequence of random variables

$$\left\{\frac{\left(\Gamma_n^{-1}\right)_{i,j}}{n^{\frac{1}{2}+\alpha}}\right\}$$

converges to zero in probability (i.p.). and some subsequence also converges with probability 1 (w.p.1) [3].

This means that for any $0 < \alpha < 1/2$.

$$\frac{1}{N^{(\frac{1}{2}+\alpha)}}\Gamma_N^{-1} \to \mathrm{diag}(N^{(\frac{1}{2}-\alpha)},\ldots,N^{(\frac{1}{2}-\alpha)}) \quad \text{i.p.}$$

and so

$$\det\left(\frac{1}{N^{(\frac{1}{2}+\alpha)}}\Gamma_N^{-1}\right) \to N^{(\frac{1}{2}-\alpha)m} \quad \text{i.p.}$$

This establishes that

$$\det(P_N) = V_N^m \det(\Gamma_N) \to (\epsilon + \delta\|f\|_1)^{2m} \quad \text{i.p.}$$

□

We now derive a deterministic asymptotic optimal input result. Consider a periodic Galois sequence having period $r = 2^n - 1$. This can be generated with n shift registers and $s < n$ (modulo two) adders, for example [11]. A periodic Galois sequence "looks" like a random binary sequence in the sense that the autocorrelation functions are very similar. Thus, one would expect a similar asymptotic result to hold in this case.

Theorem 3.3 *Let the input be a Galois sequence of period* r. *Defining* $N = qr + p$ *with* $0 < p < r$,

$$\lim_{\substack{r\to\infty \\ q\to\infty}} \sup_{g\in\mathcal{M}_0} \sup_{\|d\|_\infty\le\delta} \det(P_N) = (\epsilon + \delta\|f\|_1)^{2m}$$

Proof. We first note that $(\Gamma_N^{-1})_{i,j}$ is simply $(N-j)$ times the autocorrelation function of the input from 1 to $N-j$, evaluated at $j-i$. Now use the fact from [11] that for u, a Galois sequence of period r, and any integer $q > 0$

$$\sum_{k=1}^{qr} u_k u_{k+l} = \begin{cases} qr & l = 0 \\ -q & 1 \le l \le r-1 \end{cases}$$

Note that when $l \neq 0$, we have $\sum_{k=qr}^{(q+1)r} u_k u_{k+l} = -1$. Thus, for any $0 < p < r-1$, we have an immediate bound $|\sum_{k=qr}^{qr+p} u_k u_{k+l}| \le q + 1 + \frac{r}{2}$. The result follows after using the above facts and an arguments similar to the proof of the previous theorem which shows that the diagonal entries of Γ_N^{-1} are $\mathcal{O}(qr)$, while the off-diagonal entries are at most $\mathcal{O}(r) + \mathcal{O}(q)$. □

The above results hold for the special model structure given by Equation 17 (i.e., $W = \epsilon$). For the more general structure we can only get an upper bound for the uncertainty given optimal inputs. This is given by the following result.

Corollary 3.4 *For the model given by*

$$y = \frac{B(\theta)}{A}u + \Delta W u + d$$

$$\inf_{\|u\|_\infty \leq 1} \sup_{g \in \mathcal{M}} \sup_{\|d\|_\infty \leq \delta} \det(P_N) \leq (\|A\|_1(\|W\|_1 + \delta))^{2m} \tag{20}$$

4 Some Examples

We now consider two simple examples which illustrate the convergence of the algorithm when optimal inputs are used. The model is FIR with two parameters and is given by

$$y_k = \theta_0 u_k + \theta_1 u_{k-1} + (\Delta u)_k + d_k$$

where $\|\Delta\|_\infty \leq 0.05$, $\|d\|_\infty \leq 0.1$, and $\|u\|_\infty \leq 1$. In the first example $\theta = 0$, $\Delta = 0$, and $d = 0$. This corresponds to the worst-case situation. The input is taken as a random binary sequence in $\{-1, +1\}$. The ellipsoid is initialized to $1000I$, but shrinks significantly after the first iteration. A plot of the volume of the identified ellipsoids (for iterations 2–100) is shown in Figure 1.

In the second example, the true plant is $\theta^T = [1.0\ 0.2]$ and $\Delta = 0$. The disturbance d is chosen as a sequence of i.i.d. uniform random variables (in $[-0.1, 0.1]$) and the input is the random binary sequence. The parameter estimate is initialized to zero and the ellipsoid is initialized to $1000I$. Figure 2 shows the second, fifth, and tenth ellipsoids. The error between the estimate (ellipsoid center) and the true plant is plotted in Figure 3 while the volume is shown in Figure 4. Finally, the algorithm is run with an input which is taken from a uniform density and the convergence rate is compared in Figure 5. For iterations $\lessapprox 15$, the volumes are drastically different and cannot be compared at one scale without complete loss of detail.

5 Conclusion

In this paper we have presented a simple recursive approximate set membership identification algorithm. The model used was parametric, linear

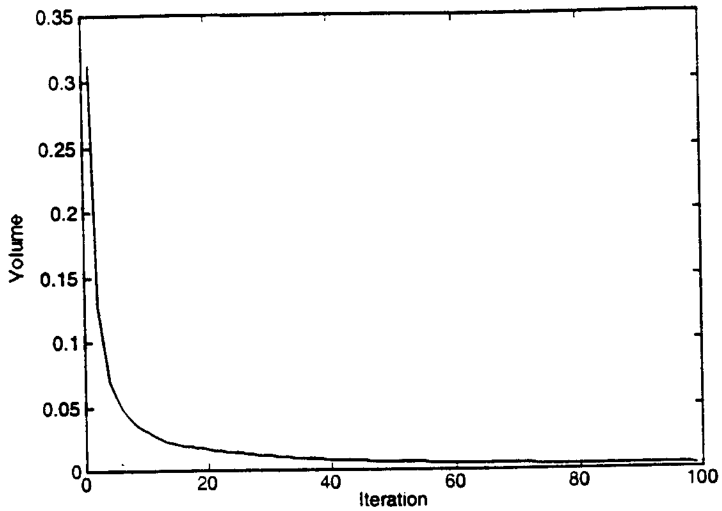


Figure 1: Worst–Case Convergence of Ellipsoid Volume

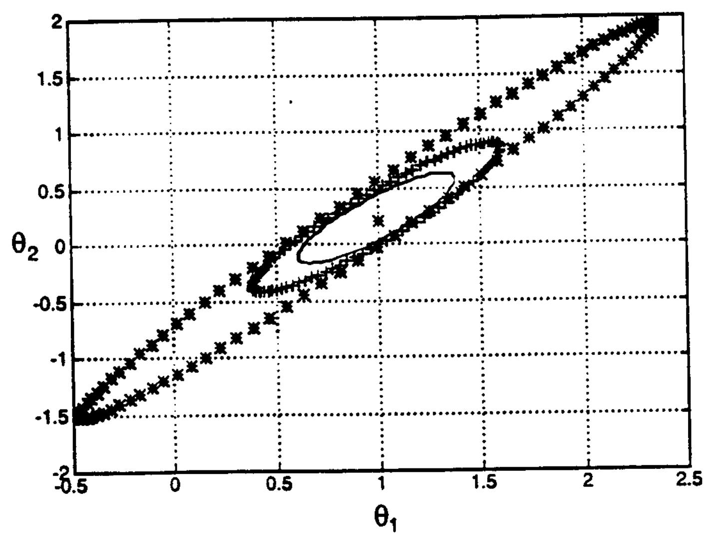


Figure 2: Evolution of Ellipsoids

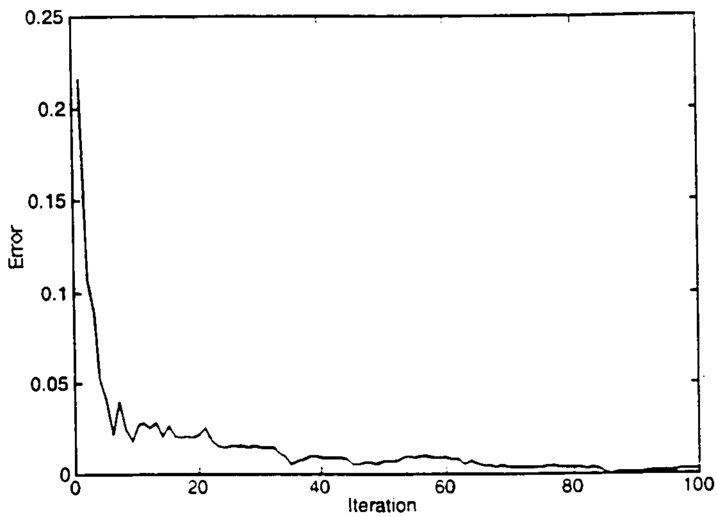


Figure 3: Convergence of Central Estimate

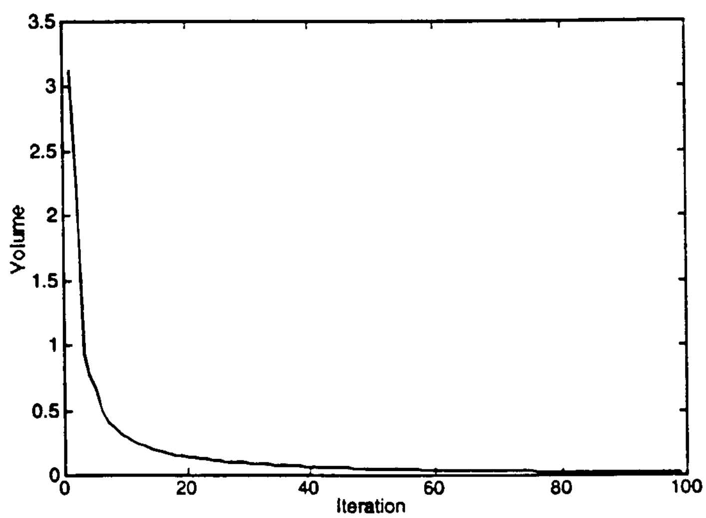


Figure 4: Optimal Convergence of Ellipsoid Volume

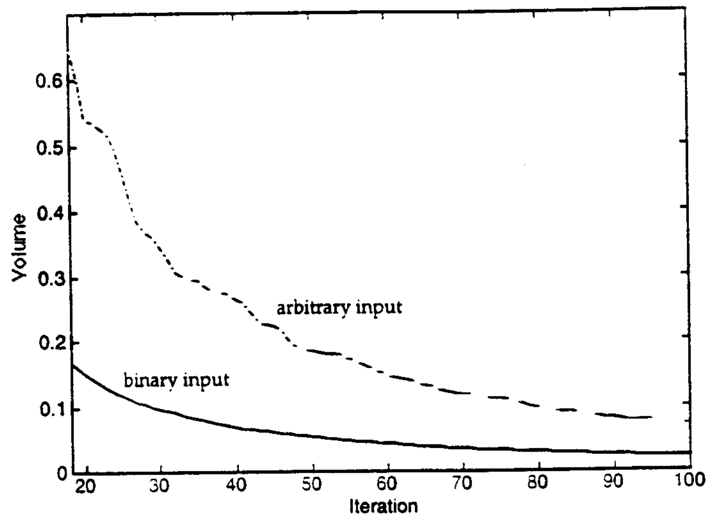


Figure 5: Convergence with Binary and Arbitrary Inputs

in parameters, and combined with nonparametric uncertainty and output additive magnitude bounded noise. The worst–case asymptotic behavior of this algorithm was studied in terms of the volume of uncertainty sets. Furthermore, it was shown that there exist inputs which can guarantee that the volume of uncertainty sets shrink to this theoretical lower bound despite worst–case conditions. A direction for future research is to extend this type of analysis to more complex model structures.

References

[1] G. Belforte, B. Bona, and V. Cerone. "Parameter Estimation Algorithms for a Set–membership Description of Uncertainty". *Automatica*, 26(5):887–898, 1990.

[2] D.P. Bertsekas and I.B. Rhodes. "Recursive State Estimation for a Set–Membership Description of Uncertainty System Inputs". *IEEE Transactions on Automatic Control*, 16:117–128, 1971.

[3] K.L. Chung. *A Course in Probability Theory*. Academic Press, 1974.

[4] J. R. Deller. 'Set Membership Identification in Digital Signal Processing". *IEEE ASSP Magazine*. pages 4–20. October 1989.

[5] J.R. Deller. M. Nayeri. and S.F. Odeh. 'Least-SquareIdentification with Error Bounds for Real-Time Signal Processing and Control". *Proceedings of the IEEE*. 81(6):815–849. June 1993.

[6] E. Fogel. 'System Identification via Membership Set Constraints with Energy Bounded Noise". *IEEE Transactions on Automatic Control*. 24(5):752–758. October 1979.

[7] E. Fogel and Y.F. Huang. 'On the Value of Information in System Identification: Bounded Noise Case". *Automatica*. 18(2):229–238. 1982.

[8] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press. 1985.

[9] R. Kosut. M. Lau. and S. Boyd. "Set-Membership Identification of Systems with Parametric and Nonparametric Uncertainty". *IEEE Trans. on Auto. Control*. 37(7):929–941, July 1992.

[10] J. Krause and P.P. Khargonekar. 'Parameter Identification in the Presence of Nonparametric Dynamic Uncertainty". *Automatica*, 26(1):113–123, 1990.

[11] E. A. Lee and D. G. Messerschmitt. *Digital Communication*. Kluwer Academic Publishers. 1988.

[12] M. Livstone. *Identification. Robust Adaptation and Iterative Schemes*. PhD thesis. MIT. Cambridge, Massachusetts. October 1994.

[13] M. Milanese and G. Belforte. "Estimation Theory and Uncertainy Intervals Evaluation in the Presence of Unknown but Bounded Errors: Linear family of Models and Estimators". *IEEE Transactions on Automatic Control*. "AC-27":408–414. 1982.

[14] J. P. Norton. "Identification and Application of Bounded Parameter Models". *Automatica*. 23(4):497–507, 1987.

[15] F.C. Schweppe. "Recursive State Estimation: Unknown but Bounded Errors and System Inputs". *IEEE Transactions on Automatic Control*. 13:22–28, 1968.

[16] R. Younce and C. Rohrs. "Identification with Non-Parametric Uncertainty". *Proc. 1990 Conference on Decision and Control. Honolulu. Hawaii.* pages 3154–3161. December 1990.

Worst-Case Identification of Nonlinear Fading Memory Systems

Munther A. Dahleh*
Lab. for Information & Decision Systems
M.I.T.
Cambridge, MA 02139

Eduardo D. Sontag‡
Dept. of Mathematics. SYCON
Rutgers University
New Brunswick, NJ 08903

David N.C. Tse†
Lab. for Information & Decision Systems
M.I.T.
Cambridge, MA 02139

John N. Tsitsiklis§
Lab. for Information & Decision Systems
M.I.T.
Cambridge, MA 02139

August 9, 1993

Abstract

In this paper, the problem of asymptotic identification for fading memory systems in the presence of bounded noise is studied. For any experiment, the worst-case error is characterized in terms of the diameter of the worst-case uncertainty set. Optimal inputs that minimize the radius of uncertainty are studied and characterized. Finally, a convergent algorithm that does not require knowledge of the noise upper bound is furnished. The algorithm is based on interpolating data with spline functions, which are shown to be well suited for identification in the presence of bounded noise; more so than other basis functions such as polynomials.

*Research Supported by AFOSR under grant AFOSR-91-0368 and by NSF under grant 9157306-ECS

†Research supported by an NSERC fellowship from the government of Canada, and by the NSF under Grant ECS-8552419

‡Research Supported by AFOSR under grant AFOSR-91-0346

§Research supported by the NSF under Grant ECS-8552419 and by AFOSR under grant AFOSR-91-0368

1 Introduction

Recently, there has been an increasing interest among the control community in the problem of identifying plants for control purposes. This generally means that the identified model should approximate the plant in the operator topology, since this allows the immediate use of robust control tools for designing controllers [2, 5]. This problem is of special importance when the data are corrupted with bounded noise. The case where the objective is to optimize prediction for a fixed input was analyzed by many researchers (see [10] and the references therein). The problem is more interesting when the objective is to approximate the original system as an operator, a problem extensively discussed in [20], especially when the plant's order is not known a priori. For linear time invariant plants, such approximation can be achieved by uniformly approximating the frequency response (H_∞-norm) or the impulse response (ℓ_1 norm). In H_∞ identification, it was shown that robustly convergent algorithms can be furnished, when the available data is in the form of a corrupted frequency response, at a set of points dense on the unit circle ([8, 6, 7]). When the topology is induced by the ℓ_1 norm, a complete study of asymptotic identification was furnished in [18] for arbitrary inputs, and the question of optimal input design was addressed. Input design has been addressed in stochastic settings (e.g. [11, 21]), but not in worst-case settings. Related work on the worst-case identification problem was also reported in [13, 14, 12, 15, 3, 9].

In this paper, the work of Tse *et al* [18] is extended to analyse the worst-case asymptotic identification of nonlinear fading memory systems. As in [18], the study is done in two steps. The first step is concerned with obtaining tight upper and lower bounds on the optimal achievable error by *any* identification algorithm. The bounds are functions of the input used for the experiments, and this can be arbitrary. The second step is then to study these bounds and characterize the inputs that will minimize them. In particular, simple topological conditions are furnished that guarantee the existence of an algorithm with a worst-case error within a factor of two from the lower bound. An *near* optimal input is characterized so that the worst-case error is within a factor of two of the bound on the noise.

It is noted that for the results on arbitrary experiments, the suggested optimal algorithms are tuned to the knowledge of the bound on the noise. If however, the near optimal input is used, then an untuned algorithm can be provided that results in a worst-case error equal to the noise bound, δ. Such an algorithm is based on interpolating data by spline functions of several variables.

The rest of the paper is organized as follows. Section 2 gives a formal definition of nonlinear fading memory systems. Section 3 describes the identification set-up. Section 4 characterizes the asymptotically optimal algorithms and the associated optimal worst-case errors for a given input. The problem of optimal inputs is addressed in Section 5. An optimal untuned algorithm is developed in Section 6. Section 7 contains our conclusions.

2 Fading Memory Systems

Let $\mathcal{U}$ be the set of one-sided infinite sequences whose ℓ_∞ norm is bounded by 1. This can be viewed as the input set which contains the possible inputs that can be used for performing the identification experiments. We consider the set of models $\mathcal{X}$ as discrete-time, causal functions from $\mathcal{U}$ to $\Re^\infty$; a plant $h \in \mathcal{X}$ takes as input a sequence $u = (u_0, u_1, \ldots)$ to give an output sequence $(h_0(u), h_1(u), \ldots)$. We assume that $h \in \mathcal{X}$ further satisfies the following properties:

1. $h_n(u)$ depends continuously on $u_0, \ldots, u_{n-1}$.

2. h has equilibrium-initial behavior:

$$h_{n+1}(0u) = h_n(u) \text{ for all } n ,$$

where $0u$ is the input $0, u_0, u_1, \ldots$. (In general, we will use the notation vw for concatenation, i.e. first apply the finite sequence v, then w. Since we are dealing with causal systems, we shall slightly abuse the notation and write $h_n(w)$ to mean $h_n(u)$, where u is any infinite sequence the first n elements of which are given by the finite sequence w.)

3. h has *fading memory* (FM): for each $\varepsilon > 0$ there is some $T = T(\varepsilon)$ such that for every k, every $t \geq T$ and every finite sequences $v \in [-1,1]^k$, $w \in [-1,1]^t$,

$$|h_{t+k}(vw) - h_t(w)| < \varepsilon$$

To measure the identification error, we shall consider the metric ρ to be the one corresponding to the operator gain:

$$\|h\|_{\mathcal{X}} = \sup_{u \in \mathcal{U}} \|h(u)\|_\infty .$$

It can be seen that systems in $\mathcal{X}$ satisfying the above property necessarily must have bounded operator gain. This is a good norm to consider for robust control applications. However, it should be noted that this norm is different from the standard definition of the gain of a nonlinear operator, which is readily suitable for robust control applications. For the above induced norm to be useful, an upper bound on the amplitudes of input signals has to be known apriori. In the above definition, this bound is normalized to one.

Examples of FM Systems:

Stable LTI systems.
For each $h \in \ell_1$ consider the input/output map $u \mapsto u * h$. It is clear that these systems satisfy the above conditions. The operator-induced norm in this case is just the ℓ_1 norm.

Hammerstein systems.
These are systems which are formed by composition of a stable LTI system followed by a memoryless nonlinear element:

$$y_n = g((u * h)_n)$$

for some $h \in \ell_1$ and some continuous function $g : \Re \to \Re$. It is easily to verify that these systems satisfy the first two conditions above. If we assume further that g is uniformly continuous, then it can be seen that the system also has fading memory.

For further details on fading memory operators, see [1, 16].

3 Identification Setup

The plant to be identified is known to be in a *model set* $\mathcal{M} \subset \mathcal{X}$. An input u is selected from the set $\mathcal{U}$. We assume that the observed output y is corrupted by some additive disturbance d which is unknown but magnitude-bounded, $\|d\|_\infty \leq \delta$, i.e. if h is the system, then

$$y = h(u) + d .$$

An *identification algorithm* is a sequence of mappings $\phi = \{\phi_n\}$ generating at each time an estimate $\phi_n(P_n u, P_n y) \in \mathcal{X}$ of the unknown plant. Here. $P_n(u_0, u_1, \ldots, u_n, u_{n+1}, \ldots) = (u_0, \ldots, u_n, 0, 0, \ldots)$ is the truncation operator.

Given an identification algorithm and a chosen input. we would like to consider the limiting situation when longer and longer of the output sequences are observed. To this end. the worst case asymptotic error. $e_\infty(\phi, u, \delta)$, of an algorithm ϕ is defined as the smallest number r such that for all plants $h \in \mathcal{M}$ and for all disturbances $\mathbf{d}$ with $\|\mathbf{d}\|_\infty \leq \delta$,

$$\limsup_{n\to\infty} \|\phi(P_n u, P_n(u * h + d)) - h\|_{\mathcal{X}} \leq r$$

Equivalently,

$$e_\infty(\phi, u, \delta) \equiv \sup_{h\in\mathcal{M}} \sup_{\|\mathbf{d}\|_\infty \leq \delta} \limsup_{n\to\infty} \|\phi(P_n u, P_n(u * h + d)) - h\|_{\mathcal{X}}$$

The interpretation of this definition is that no matter what the true plant and the disturbances are, the plant can be eventually approximated to within $e_\infty(\phi, u, \delta)$, using the estimates generated by the identification algorithm. The convergence rate may depend on the plant and noise. i.e. for a given ε there exists some $N(d, h, \varepsilon)$ so that

$$\|\phi_n(y) - h\|_{\mathcal{X}} \ < \ e_\infty(\varphi, u, \delta) + \varepsilon$$

whenever $n \geq N$. We say that the convergence is uniform if $N(y, h, \varepsilon)$ depends only on ε. For more motivations and discussions on these definitions. see [18].

The optimal worst-case asymptotic error $E_\infty(u, \delta)$ is defined as the smallest error achievable by any algorithm:

$$E_\infty(u, \delta) \equiv \inf_\phi e_\infty(\phi, u, \delta)$$

Any algorithm for which the infimum is attained is said to be *asymptotically optimal.* We will obtain a general characterization of the asympotically optimal algorithms and the resulting optimal error, for any given input u. We will then find conditions on the input u to make this optimal worst-case asymptotic error small.

4 Asymptotically Optimal Identification

The characterization of asymptotically optimal algorithms and optimal asymptotic errors is in terms of the *uncertainty set*, an important notion in information-based complexity theory. The uncertainty set $S_n(u, y, \delta)$ at time n is the set of all systems in the given model set $\mathcal{M}$ which are consistent with the observed data up until time n:

$$S_n(u, y, \delta) = \{h \in \mathcal{M} : \|P_n(y - h(u))\|_\infty \leq \delta\}$$

These are the plants which can give rise to the observed output for some valid disturbance sequence. The infinite-horizon uncertainty set is

$$S_\infty(u, y, \delta) = \{h \in \mathcal{M} : \|y - h(u)\|_\infty \leq \delta\}$$

For a given set $A \subset \mathcal{X}$, define the diameter of the set as:

$$\operatorname{diam}(A) = \sup_{g,h\in A} \|g - h\|_{\mathcal{X}}$$

and let $D(u,\delta)$ be the diameter of the worst-case infinite-horizon uncertainty set:

$$D(u,\delta) = \sup_{h\in\mathcal{M}} \sup_{\|d\|_\infty \leq \delta} \operatorname{diam}(S_\infty(u, u*h+d,\delta))$$

Under appropriate topological conditions on the model set, this quantity characterizes the optimal asymptotic worst-case error. The following result is a generalization of Proposition 3.3, Theorem 3.4 and Proposition 3.9 in the LTI case ([18]) to our present setting.

Theorem 4.1 *If the model set $\mathcal{M} \subset \mathcal{X}$ is σ-compact (i.e. $\mathcal{M}$ is a countable union of compact sets), then*

$$\frac{D(u,\delta)}{2} \leq E_\infty(u,\delta) \leq D(u,\delta)$$

Furthermore, if $\mathcal{M}$ is compact, then the convergence can be made uniform.

In the σ-compact case, an algorithm achieving an asymptotic error within the above bounds can be realized using the principle of *Occam's razor*. Let $\mathcal{M} = \cup_i \mathcal{M}_i$, where the $\mathcal{M}_i$'s are compact and increasing. This decomposition gives a complexity index to each plant in $\mathcal{M}$, as the index of the smallest M_i containing the plant. At each time n, the algorithm simply returns as an estimate any plant in the uncertainty set S_n with the smallest complexity index. Note that since the disturbance bound δ is required to compute the uncertainty set, this algorithm is tuned to this information. On the other hand, if $\mathcal{M}$ is compact, one can use an algorithm which simply returns the plant in $\mathcal{M}$ which fits best the input/output data observed so far. This algorithm attains an asymptotic error within the above bounds with a uniform rate of convergence. It is also untuned to the disturbance bound δ.

A slight extension of the above result yields essentially the same bounds for the case when $\mathcal{M}$ is separable. The proof is along the same lines as the proofs of Lemma 4.5 and Proposition 4.6 in [18]. The optimal algorithm has roughly the same structure as that for the σ-compact case.

Theorem 4.2 *If $\mathcal{M}$ is separable, then*

$$\frac{D(u,\delta)}{2} \leq E_\infty(u,\delta) \leq \lim_{x\downarrow\delta} D(u,x)$$

To apply the above results, we now look at the topological structure of some classes of fading memory systems under the operator-induced norm.

Consider first the class of stable LTI systems. Since this corresponds to the space ℓ_1, which is separable, Theorem 4.2 is applicable in this case. More generally, we can in fact prove:

Theorem 4.3 *The class of all fading memory systems is separable.*

Proof. Define the class of pth-order memory systems, $\mathcal{M}_p$, to be the set of all f such that for every k and for every $t > p$ and every finite sequences $v \in [-1,1]^k, w \in [-1,1]^t$, $f_{t+k}(vw) = f_t(w)$. It is clear that any fading memory system can be approximated (in the operator-induced norm) arbitrarily closely by a pth-order memory system for sufficiently large p. Hence it suffices to prove that $\mathcal{M}_p$ is separable for all p.

Now given any $f \in \mathcal{M}_p$ we can find some continuous function $g : [-1,1]^p \to \Re$ such that for all time n, and all input u,

$$f_n(u) = g(u_{n-p},\ldots,u_{n-1})$$

We call g the *memory function* for f. Hence we have that $\|f\| = \|g\|_\infty$, where the infinity norm is taken over $[-1,1]^p$. But the space of continuous functions with the uniform topology induced by the ℓ_∞-norm, denoted by $C([-1,1]^p)$, is separable, and hence so is $\mathcal{M}_p$. ∎

This means that when we look at fading memory system, we can apply Theorem 4.2, and reduce the analysis of the asymptotic optimal error to the analysis of the worst-case infinite-horizon diameter.

5 Optimal Inputs

We now turn to the question of optimal inputs, i.e. inputs u that minimize the worst-case infinite-horizon diameter $D(u,\delta)$. First we state a simple lower bound. Let

$$\rho(\mathcal{M}) := \sup\{r \mid 0 < \delta < r \Rightarrow \exists\, g, h \in \mathcal{M} \text{ with } \|g - h\| = r\}\,.$$

Note that if $\mathcal{M}$ is path connected, then $\rho(\mathcal{M}) = \mathrm{diam}(\mathcal{M})$.

Lemma 5.1 If $2\delta < \rho(\mathcal{M})$, then $D(u,\delta) \geq 2\delta$ for all $u \in \mathcal{U}$.

Proof. See [18] ∎

Since $\rho(\mathcal{M}) > 2\delta$ for most of the reasonable model sets, the above result gives a general lower bound. We now investigate how to choose an input which achieves this bound.

Recall that $\mathcal{M}$ is *balanced* if $h \in \mathcal{M}$ implies $-h \in \mathcal{M}$. For balanced and convex model sets, it is well known from information-based complexity theory [17] that the worst case diameter is equal to the diameter of the uncertainty set when the output is identically equal to zero. The following lemma summarizes this.

Lemma 5.2 Assume that $\mathcal{M}$ is balanced and convex. Then, for all $u \in \mathcal{U}, \delta > 0$,

$$D(u,\delta) = \mathrm{diam}(S_\infty(u,0,\delta))$$

Call an input $u \in \mathcal{U}$ *persistently exciting* for $\mathcal{M}$ if the following property holds:

$$\|h(u)\|_\infty = \|h\|_{\mathcal{X}} \quad \forall\, h \in \mathcal{M}\,.$$

The following result says that persistently exciting inputs are optimal.

Theorem 5.1 *Assume $\mathcal{M}$ is balanced and convex.*

1. *If the input u is persistently exciting, then $D(u,\delta) \leq 2\delta$ for all $\delta > 0$.*

2. *If u is persistently exciting then $D(u,\delta) = 2\delta$ for each $0 < \delta \leq \frac{\mathrm{diam}(\mathcal{M})}{2}$.*

Proof. (1): By Lemma 5.2, for all $\delta > 0$,

$$D(u,\delta) \;=\; 2\sup\{\|h\| \mid h \in \mathcal{M}, \|h(u)\|_\infty \leq \delta\}\,.$$

Pick any $h \in \mathcal{M}$ such that $\|h(u)\|_\infty \leq \delta$. If u is persistently exciting, this means that also $\|h\|_{\mathcal{X}} \leq \delta$, so $D(u,\delta) \leq 2\delta$.

(2) From Lemma 5.1, $D(u,\delta) \geq 2\delta$ for such δ. The result follows from (1) above. ∎

It follows from Theorem 4.2, Theorem 4.3, Lemma 5.1 and the above theorem that one can achieve nearly optimal asymptotic identification for the entire class of fading memory systems if we use a persistently exciting input.

Corollary 5.3 Let $\mathcal{M} = \mathcal{X}$, the class of all fading memory systems. Then for any identification algorithm ϕ and any input u, the worst-case asymptotic error $e_\infty(\phi, u, \delta)$ is lower bounded by δ. If u is persistently exciting, then there is an algorithm which can achieve an asymptotic error of less than 2δ.

A natural question which arises at this point is whether persistently exciting inputs exist. In the stable LTI case, this was shown to be the case [18]. The next theorem shows that they also exist when the model set consists of nonlinear fading memory systems.

Theorem 5.2 *Let the model set $\mathcal{M}$ be some subset of the set of fading memory systems. Let W be any countable dense subset of $[-1,1]$ and consider any input $\omega_0 \in [-1,1]^\infty$ which contains all possible finite sequences of elements of W. Then, ω_0 is persistently exciting.*

Proof. Assume that $h \in \mathcal{M}, \|h\| = K < \infty$. Pick any $\varepsilon > 0$. Let $T = T(\varepsilon)$ as in the definition of FM. By definition of the sup norm, there is some ω and some T_1 so that

$$\sup_{0 \leq t < T_1} |h_t(\omega)| > K - \varepsilon .$$

Using the equilibrium-initial assumption and replacing ω by $0^T\omega$ and T_1 by $T + T_1$, we may assume that

$$\sup_{T \leq t < T_1} |h_t(\omega)| > K - \varepsilon .$$

By density of W and continuity of $h_t(\omega)$ on past values of ω, we may further assume that $\omega(0), \ldots, \omega(T_1 - 1)$ take values in W. From the construction of ω_0, there is some k so that

$$\omega_0(k) = \omega(0),\ \omega_0(k+1) = \omega(1),\ \ldots,\ \omega_0(k + T_1 - 1) = \omega(T_1 - 1) .$$

Let v be the finite sequence $\omega_0(0), \omega_0(1), \ldots, \omega_0(k-1)$ and w be the finite sequence

$$\omega(0), \omega(1), \ldots \omega(T_1 - 1) ,$$

which is equal to $\omega_0(k), \omega_0(k+1), \ldots, \omega_0(k + T_1 - 1)$. So vw is the same as the first $T_1 + k - 1$ elements of ω_0.

By the FM property applied to these inputs, we have that

$$|h_{t+k}(vw) - h_t(w)| < \varepsilon \text{ for each } t \geq T$$

(using the notational convention mentioned above for $h_s(w)$ if the length of w is larger than s). Then for such t,

$$|h_{t+k}(\omega_0)| = |h_{t+k}(vw)| \geq |h_t(w)| - \varepsilon ,$$

so

$$\|h(\omega_0)\| \geq \sup_{T+k \leq r < T_1 + k} |h_r(\omega_0)| \geq K - 2\varepsilon .$$

Thus, we conclude that $K = \|h\| \geq \|h(\omega_0)\| \geq K - 2\varepsilon$ for all $\varepsilon > 0$, so $\|h(\omega_0)\| = K$ as wanted. ∎

6 An Untuned Algorithm

As remarked earlier, the asymptotically optimal algorithms for σ-compact and *separable* model sets are tuned to the knowledge of δ, the bound on the disturbance. It will be shown that for fading memory systems, one can achieve asymptotically optimal identification without knowing δ, provided that we use a persistently exciting input. This is in fact a generalization of a result by Makila [13], which was proved in the context of stable LTI systems.

We shall make use of multivariate piecewise linear spline functions to interpolate between the measured data to form an approximation to the unknown plant. This is a generalization of the univariate linear spline, but because in higher dimension there is no natural ordering of the data points, the description of the interpolant is more complicated.

Consider the cube $I = [-1,1]^p \subset \Re^p$. Let $x_1, x_2, \ldots, x_m, m > p$ be m points in the interior of the cube. We wish to construct a continuous, piecewise linear function $f : I \rightarrow \Re$ such that $f(x_i) = y_i, i = 1, 2, \ldots, m$, where the y_i's are given data values to interpolate.

To facilitate the discussion, we need to first define several basic geometric concepts. A p-dimensional *simplex* S in $\Re^p$ is the convex hull of $p+1$ affinely independent points. Each of these points is a *vertex* of S. The convex hull F of any subset of these $p+1$ points form a *face* of S if there exists a hyperplane H such that S lies entirely on one side of H and $S \cap H = F$. If F is the convex hull of p points, it is called a *facet*. A point v outside S is said to be separated from S by a face F if v and S lie on the opposite sides of the $p-1$ dimensional hyperplane containing F.

The first step is to find a set of simplices $\{S_j\}$ such that (1) their combined vertex set is $\{x_1, \ldots, x_m\}$, (2) the simplices only intersect at common faces (3) their union gives the convex hull of the vertex set. This can be done inductively as follows: for $m = p+1$, the set simply consists of one simplex which is the convex hull of the $p+1$ points. Suppose now we obtain a set of simplices $S_1, S_2, \ldots, S_d$ to cover $m > p$ points, and consider one additional point x_{m+1}. If $x_{m+1} \in S_k$ for some k, then we can simply replace S_k with the $p+1$ simplices formed by x_{m+1} with each of the faces of S_k. It is easy to see that these $p+1$ simplices only intersect at common faces and their union is S_k, so that the updated set of simplices now covers the $m+1$ points. On the other hand, if x_{m+1} lies outside $P = \cup_{i=1}^{d} S_i$, then for each facet F of some S_k which separates x_{m+1} from P, we add a simplex formed by x_{m+1} with F to the set. It can also be proved that these added simplices together with the original ones satisfy the three conditions.

Given these simplices, we can now define our interpolating linear spline f as follows. First define $f(x_i) = y_i$ at the given data points. For other $x \in [-1,1]^p$, if $x \in S_j$ for some j, let $f(x) \equiv \sum_i \alpha_i f(u_i)$, where u_i's are the vertices of S_j and $x = \sum_i \alpha_i u_i$. It is easy to check that because of the above three conditions on the simplices, f is well-defined and continuous. To extend f continuously outside $P = \cup_j S_j$, define $f(x)$ to be equal to the value of f at the nearest point in P to x. Since P is convex, this nearest point is unique, and this guarantees the continuity of this extension.

If we view this interpolating process as an operator T_m mapping the data vector $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ to the piecewise linear interpolating function $(T_m(\mathbf{y}))(x)$, then we can see that this operator is linear and its gain defined as:

$$\|T_m\| \equiv \sup_{\|\mathbf{y}\|_\infty = 1} \|T_m(\mathbf{y})\|_\infty \tag{1}$$

is equal to one. This simple fact ensures that no matter how many data are obtained, noise in the data will not be amplified in the interpolating process. This property of linear splines, which is not shared by methods such as global polynomial interpolation, turns out to be the key to guarantee

the consistency of the estimates. A similar situation is encountered in linear system identification from frequency response data [8], where 1 dimensional splines are used instead of polynomials to interpolate the noisy data to guarantee robustness of the identification procedure.

With the above basic discussions on multivariate linear splines, we may now state the main result of this section.

Theorem 6.1 *Let the model set $\mathcal{M} = \mathcal{X}$, the set of all fading memory systems. If the input u is persistently exciting, then there is an algorithm which can achieve an optimal worst-case asymptotic error $e_\infty(\phi, u, \delta) = \delta$. This algorithm does not require the knowledge of δ in computing the estimates.*

Proof. The structure of the algorithm is as follows. We view the model set $\mathcal{M}$ as the closure of the finite-memory systems $\mathcal{M}_p$, $p = 0, 1, 2, \ldots$. We start by assuming that the true system is in $\mathcal{M}_0$. Data is observed until time $n(0)$, after which the algorithm comes up with an estimate $\hat{h}^{(0)} \in \mathcal{M}_0$. Then it moves onto the next model set $\mathcal{M}_1$, and waits until time $n(1)$ before coming up with an estimate $\hat{h}^{(1)} \in \mathcal{M}_1$. The algorithm continues to move onto the model set of one higher order, to produce a new estimate. It will be shown below how the time $n(p)$ is specified and the estimate $\hat{h}^{(p)}$ is computed for each p.

Let h be the true system. Let $\{\delta_p\}$ be any sequence which monotonically goes to zero.

Fix p, and let the time $n \in [n(p-1), n(p)]$. (This is when the algorithm is collecting data to compute an estimate in $\mathcal{M}_p$.) Consider all the blocks

$$(u_{n-p+1}, \ldots, u_{n-1}, u_n), \ \forall n = n(p-1), \ldots, n(p)$$

in the input as data points in the cube $[-1, 1]^p$. We maintain a simplex structure in $[-1, 1]^p$ with these data points as vertex set, and the structure is incrementally modified more or less according to the procedure discussed earlier, with a slight twist. Let $C_n = \cup_j S_j$ be the union of the simplices at time n, and d_n be the distance between C_n and the corner of $[-1, 1]^p$ farthest away from C_n. At time $n + 1$, one more data point is obtained. If $d_n < \delta_p$ and the new data point lies outside C_n, then discard the new point. Otherwise update the simplex structure as described earlier.

Let $n(p)$ be the earliest time such that $d_{n(p)} < \delta_p$ and the diameter of the largest simplex in C_n is less than δ_p. At this time, the algorithm returns an estimate $\hat{h}^{(p)} \equiv \phi_{n(p)}(h(u) + d)$ to be the p-th order system with memory function as the piecewise linear spline interpolant of the current simplex structure.

We now claim that $n(p) < \infty$ for every p. First we see that because the input is persistently exciting, the p-blocks in u are dense in $[-1, 1]^p$ (Otherwise, there is a ball in $[-1, 1]^p$ which does not contain any blocks in u, and we can construct a p-step finite memory system with a continuous memory function $f : \Re^p \to \Re$ to be positive at the centre of the ball and zero outside. Then applying the input u to the system will give a zero output while the gain of the system is non-zero, thus contradicting the persistent excitedness of u.) Hence there exists a time $m(p)$ such that $d_{m(p)} < \delta_p$. After this time, the convex hull C_n no longer expands. All the changes consist of further partitioning of the simplices inside C_n due to the new data points. Because the data points are dense, it can seen that the diameter of the largest simplex must go to zero. Hence, $n(p)$ is finite.

We now claim that:

$$\limsup_{p \to \infty} \|\phi_{n(p)}(h(u) + d) - h\| \leq \delta$$

for all d, $\|d\|_\infty \leq \delta$. Note also that the algorithm defined above does not use the value of δ in computing the estimate.

Take any $\epsilon > 0$. There exists some q such that $\mathcal{M}_q$ contains a system h^ϵ with

$$\|h - h^\epsilon\| \leq \epsilon \tag{2}$$

Let g^ϵ be the (q-step) memory function for h^ϵ. For $p \geq q$, $\phi_{n(p)}(h(u)+d)$ is the spline interpolant that approximates the unknown memory function, and $y = h(u) + d$ is the output. We can also extend g^ϵ to a function on $[-1, 1]^p$ which depends only on the last q coordinates. Now,

$$\begin{aligned}
& \|\phi_{n(p)}(h(u) + d) - g^\epsilon\|_\infty \\
= \; & \|\phi_{n(p)}(h(u)) + \phi_{n(p)}(d) - g^\epsilon\|_\infty \qquad \text{by linearity of the interpolation operator} \\
\leq \; & \|\phi_{n(p)}(h(u)) - g^\epsilon\|_\infty + \|\phi_{n(p)}(d)\|_\infty \\
\leq \; & \|\phi_{n(p)}(h(u)) - g^\epsilon\|_\infty + \delta \qquad \text{by eq. 1.}
\end{aligned}$$

The first term is the interpolation error when the data is noiseless, whereas the second term is the error due to the noise. We now show that the first term can be made arbitrarily small for large p.

Since g^ϵ is continuous, g^ϵ is a uniformly continuous function on $[-1, 1]^q$. Choose ϵ_1 such that

$$\|x_1 - x_2\|_2 \leq \epsilon_1 \Rightarrow \|g^\epsilon(x_1) - g^\epsilon(x_2)\|_2 \leq \epsilon . \tag{3}$$

Now pick p sufficiently large such that $\delta_p < \epsilon_1$ and $p > q$. Let $g^p(x) \equiv \phi_{n(p)}(h(u))$.

Now for any $x \in C_{n(p)}$, the convex hull, let $x = \sum_i \alpha_i x_i$, where x_i are the vertices of the simplex containing x. Since g^p agreees with the noiseless output data at the vertices, by Eqn. 2, for each i,

$$|g^p(x_i) - g^\epsilon(x_i)| \leq \epsilon . \tag{4}$$

We have:

$$\begin{aligned}
& |g^p(x) - g^\epsilon(x)| \\
= \; & |\sum_i \alpha_i g^p(x_i) - g^\epsilon(x)| \\
\leq \; & |\sum_i \alpha_i g^\epsilon(x_i) - g^\epsilon(x)| + \epsilon \qquad \text{by Eqn. 4} \\
\leq \; & \sum_i \alpha_i |g^\epsilon(x_i) - g^\epsilon(x)| + \epsilon \\
\leq \; & 2\epsilon
\end{aligned}$$

by Eqn. 2, since $\|x_i - x\|$ is less than the diameter of the simplex, which is smaller than ϵ_1.

Now for x outside $C_{n(p)}$, let x' be the point in $C_{n(p)}$ which is closest to x. By definition of $n(p)$, the distance of x' from x is at most $\delta_p < \epsilon_1$. Hence:

$$\begin{aligned}
& |g^p(x) - g^\epsilon(x)| \\
= \; & |g^p(x') - g^\epsilon(x)| \qquad \text{by definition of the interpolant} \\
\leq \; & |g^p(x') - g^\epsilon(x')| + \epsilon \qquad \text{by Eqn. 2} \\
\leq \; & 3\epsilon \qquad \text{from above.}
\end{aligned}$$

Therefore, if h^p is the finite-memory system with memory function $\phi_{n(p)}(h(u)+d)$, then

$$\begin{aligned} & \|h^p - h\| \\ \leq\ & \|g^p - g^\epsilon\|_\infty + \epsilon \\ \leq\ & \delta + 4\epsilon\,. \end{aligned}$$

Since this is true for all ϵ, it follows that

$$\limsup_{p\to\infty} \|h^p - h\| \leq \delta$$

as desired. ∎

We would like to make a comment about the time complexity of this identification problem. It can be easily seen that in general, the time needed to identify a system to a prescribed accuracy grows exponentially as the order of the system, even when there is no noise. For example, if we assume a certain Lipschitz condition on the order p memory function g, such as $|g(x) - g(y)| < M\|x - y\|$, then to identify the function up to accuracy ϵ (in the $\|\cdot\|_\infty$ norm), the number of data points needed is at least the minimum number of ϵ-balls to cover $[-1,1]^p$. Since the volume of an ϵ-ball is proportional to ϵ^p, it is clear that this minimum number is at least proportional to $(\frac{1}{\epsilon})^p$, and hence so is the experiment length. This means that if p is large, the experiment length will be very long if we make no further assumption on the unknown plant.

It is interesting to compare this situation with the problem of identifying *linear* finite impulse response systems. For nonlinear systems the time complexity is exponential of the order, whether or not there is noise. For the linear case, while it takes only linear time to identify a FIR system *exactly* when there is no noise, it has been shown [3, 15] that the time complexity immediately becomes exponential once we introduce any worst-case noise. Moreover, it has been demonstrated that if we are willing to put a probability distribution on the noise, polynomial time complexity can often be obtained [19]. These facts show that while in the nonlinear case, the plant uncertainty determines the time complexity of the identification, in the linear case, the complexity is sensitive to how the noise is modelled.

7 Conclusions

A framework for the analysis of asymptotic worst-case identification of LTI systems has been extended to the setting of nonlinear fading memory systems. For model sets that are either σ-compact or separable, and for any experiment, the optimal worst-case error is always bounded by twice the lower bound, which is the diameter of a certain uncertainty set. Optimal inputs which minimize this diameter are characterized. It is also shown that accurate asymptotic identification can be achieved by an optimal input, using an untuned algorithm based on spline interpolation.

References

[1] S. Boyd and L.O. Chua, "Fading Memory and the problem of approximating nonlinear operators with Volterra series," *IEEE Transactions on Circuits and Systems*, Vol. CAS-32, 1985, pp. 1150–1161.

[2] M.A. Dahleh and M.H. Khammash, "Controller Design in the Presence of Structured Uncertainty," to appear in *Automatica* special issue on robust control.

[3]

[4] M.A.Dahleh, T.V.Theodosopoulos, J.N. Tsitsiklis " The Sample Complexity of Worst-Case Identification of FIR Linear Systems", *Systems and Control Letters*, March. 1993.

[5] J. C. Doyle. "Analysis of feedback systems with structured uncertainty," *IEEE Proceedings* 129, 242-250,1982.

[6] G. Gu, P.P. Khargonekar and Y. Li, "Robust convergence of two stage nonlinear algorithms for identification in $\mathcal{H}_\infty$, *Systems and Control Letters*, Vol 18, No. 4, April 1992.

[7] G. Gu and P.P. Khargonekar, " Linear and nonlinear algorithms for identification in $\mathcal{H}_\infty$ with error bounds, *IEEE Trans. A-C*, Vol 37, No. 7, July 1992, pp 953-963.

[8] A.J. Helmicki, C.A. Jacobson and C.N. Nett, "Control-oriented System Identification: A Worst-case/deterministic Approach in H_∞'," *IEEE Trans. A-C*, Vol 36, No. 10, October 1991, pp 1163-1176.

[9] C.A. Jacobson and C.N. Nett, "Worst-case system identification in ℓ_1: Optimal algorithms and error bounds," in *Proc. of the 1991 American Control Conference*, June 1991.

[10] L. Ljung, *System Identification - Theory for the User*, Prentice-Hall, 1987.

[11] R.K. Mehra, "Optimal input signals for parameter estimation in dynamic systems–A survey and new results", em IEEE Trans. Automatic Control, vol AC-19, pp.753-768, 1974.

[12] B. Kacewicz and M. Milanese, "On the Optimal Experiment Design in the Worst-Case ℓ_1 System Identification, submitted to *CDC*.

[13] P.M. Makila, "Robust Identification and Galois Sequences", Technical Report 91-1, Process Control Laboratory, Swedish University of Abo, January, 1991.

[14] P.M. Makila and J.R. Partington, "Robust Approximation and Identification in H_∞", Proc. 1991 American Control Conference, June, 1991.

[15] K. Poolla and A. Tikku, "On the time complexity of worst-case system identification," preprint, 1992.

[16] J.S. Shamma and R. Zhao, "Fading memory feedback systems and robust stability," to appear in *Automatica*.

[17] J.F. Traub and H. Wozniakowski, *A General Theory of Optimal Algorithms*, Academic Press, New York, 1980.

[18] D. Tse, M.A. Dahleh and J.N. Tsitsiklis. Optimal Asymptotic Identification under bounded disturbances. To appear in the *IEEE Trans. Automat. Contr.*

[19] D.N.C. Tse and J.N. Tsitsiklis, " Sample Complexity of Worst-Case System Identification in the Presence of Bounded Random Noise", in preparation.

[20] G. Zames, "On the metric complexity of casual linear systems: ϵ-entropy and ϵ-dimension for continuous-time", *IEEE Trans. on Automatic Control*, Vol. 24, April 1979.

[21] M. Zarrop, *Optimal Experimental Design for Dynamic System Identification*, Springer-Verlag, New York, 1979.

T-AC/38/5//08453

Minimization of the L^{∞}-Induced Norm for Sampled-Data Systems

Bassam Bamieh
Munther A. Dahleh
J. Boyd Pearson

Reprinted from
IEEE TRANSACTIONS ON AUTOMATIC CONTROL
Vol. 38, No. 5, May 1993

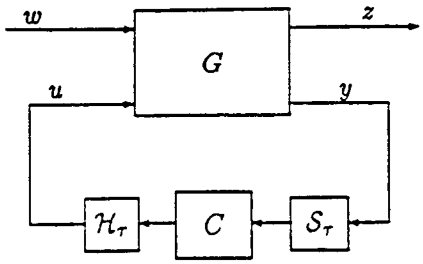


Fig. 1. Hybrid discrete/continuous-time system.

inter-sample dynamics of the hybrid system, and that the inter-sample dynamics are governed only by the plant and not the controller dynamics. We use the latter fact to derive explicit bounds on the approximation [main inequality (5)] which can be computed *a priori* and depend only on the plant. We also show that the rate of convergence of the approximation is $(1/n)$.

As already mentioned, sampled-data systems are periodic, the main theoretical tool we use for dealing with periodic systems is a *lifting* technique for continuous-time systems developed in [1], [2].[1] The technique establishes a strong correspondence between periodic systems and time invariant infinite-dimensional systems. In the next section, we briefly describe the lifting and it's application to the sampled-data problem. We then set up an equivalent infinite-dimensional problem whose solution is obtained using an approximation procedure. Formulas for the (almost) equivalent discrete-time problem are given in Section III. In the later sections, the issue of the convergence of the approximation procedure is investigated, this is done by decomposing the equivalent infinite-dimensional problem and analyzing the decomposition. In the last section, a geometric interpretation is given for the reduction of the infinite-dimensional problem, and it is compared with the $\mathscr{H}^\infty$ sampled-data problem from [1]. We also discuss possible reasons behind the fact that in the l^1 sampled-data problem (in contrast to the $\mathscr{H}^\infty$ sampled-data problem), the solutions are given by approximation, rather than exact procedures.

Finally, we note that although the closed loop, sampled-data system is periodically time varying, and thus one cannot refer to the l^1 norm of its impulse response, it is shown in [3] that the L^∞-induced norm of a periodic system can be interpreted as a type of an l^1 norm of the operator-valued "impulse response" of the lifted system. This justifies calling this problem the l^1 sampled-data problem.

II. The Lifting Technique in Sampled-Data Systems

In this section, we briefly summarize the lifting technique for continuous-time periodic systems developed in [1], [2], and apply it to the sampled-data problem. The idea of the lifting technique is to put a periodic continuous-time system in a strong correspondence with a shift invariant (i.e., discrete-time time-invariant) system, which amounts to rearranging the original system so that its periodicity can be viewed as shift invariance. To accomplish this, we first define the lifting for signals, for which the appropriate signal spaces need to be established.

For continuous-time signals, we consider the usual $L^\infty[0,\infty)$ space of essentially bounded functions [8], and it's extended version $L_e^\infty[0,\infty)$. We will also need to consider discrete-time signals that take values in a function space, for this, we define l_X to be the space of all X-valued sequences, where X is some Banach space. We define l_X^∞ as the subspace of l_X with bounded norm sequences, i.e., where for $\{f_i\} \in l_X$, the norm $\|\{f_i\}\|_{l_X^\infty} := \sup_i \|f_i\|_X < \infty$. Given any $f \in L_e^\infty[0,\infty)$, we define its *lifting* $\hat{f} \in l_{L^\infty[0,\tau]}$, as follows: $\hat{f}$ is an $L^\infty[0,\tau]$-valued sequence, we denote it by $\{\hat{f}_i\}$, and for each i

$$\hat{f}_i(t) := f(t+\tau i) \qquad 0 \le t \le \tau.$$

The lifting can be visualized as taking a continuous-time signal and breaking it up into a sequence of "pieces" each corresponding to the function over an interval of length τ (see Fig. 2). Let us denote this lifting by W_τ: $L_e^\infty[0,\infty) \to l_{L^\infty[0,\tau]}$. W_τ is a linear isomorphism, furthermore, if restricted to $L^\infty(0,\infty)$, then W_τ: $L^\infty[0,\infty) \to l^\infty_{L^\infty[0,\tau]}$ is an isometry, i.e., it preserves norms.

Using the lifting of signals, one can define a lifting on systems. Let G be a linear continuous-time system on $L_e^\infty[0,\infty)$, then its *lifting* $\hat{G}$ is the discrete-time system $\hat{G} := W_\tau G W_\tau^{-1}$, this is illustrated in the commutative diagram below:

$$\begin{tikzcd} l_{L^\infty[0,\tau]} \arrow[r, "\hat{G}"] \arrow[d, "W_\tau^{-1}"'] & l_{L^\infty[0,\tau]} \\ L_e^\infty[0,\infty) \arrow[r, "G"] & L_e^\infty[0,\infty) \arrow[u, "W_\tau"'] \end{tikzcd}.$$

Thus, $\hat{G}$ is a system that operates on Banach space ($L^\infty[0,\tau]$) valued signals, we will call such systems infinite dimensional. Note that since W_τ is an isometry, if G is stable, i.e., a bounded linear map on L^∞ then $\hat{G}$ is also stable, and furthermore, their respective induced norms are equal, $\|\hat{G}\| = \|G\|$. The correspondence between a system and its lifting also preserves algebraic system properties such as addition, cascade decomposition and feedback (see [1] for details).

The usefulness of the lifting in the sampled-data problem is the fact that if G is a τ-periodic system, then $\hat{G}$ commutes with the shift on $l_{L^\infty[0,\tau]}$, that is, $\hat{G}$ is shift invariant. This basic fact allows us to treat continuous-time periodic systems as discrete-time time-invariant systems, albeit infinite-dimensional systems.

State space models can be found for the lifted systems. To illustrate, let G be a continuous-time time-invariant system given by a state space realization $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$. In [1] it was shown that the lifting $\hat{G}$ has a state space

[1] Essentially the same technique was arrived at independently in [22] and [23].

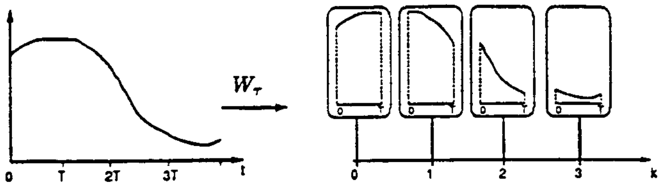


Fig. 2. W_τ: $L_e^p[0,\infty) \rightarrow l_{L^p[0,\tau]}$.

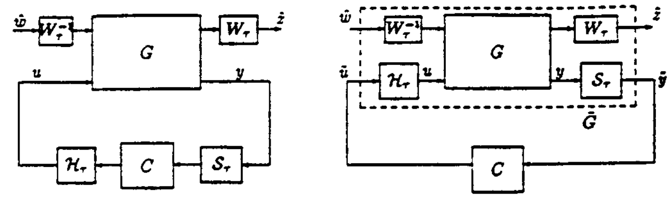


Fig. 3. Equivalent problem.

realization given by:

$$\hat{G} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] = \left[\begin{array}{c|c} e^{A\tau} & e^{A(\tau-\hat{s})}B \\ \hline Ce^{A\hat{t}} & Ce^{A(\hat{t}-\hat{s})}1_{(\hat{t}-\hat{s})}B + D\delta(\hat{t}-\hat{s}) \end{array}\right]$$

$$\begin{aligned} \hat{B}&: \quad L^\infty[0,\tau] \rightarrow \mathbb{R}^{n_x} \\ \hat{A}&: \quad \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x} \\ \hat{C}&: \quad \mathbb{R}^{n_x} \rightarrow L^\infty[0,\tau] \\ \hat{D}&: \quad L^\infty[0,\tau] \rightarrow L^\infty[0,\tau] \end{aligned} \tag{1}$$

where the operators $\hat{C}, \hat{B}, \hat{D}$ are given in terms of their kernel functions, and $1_{(\cdot)}$ is the unit step function.

Notation: It simplifies the notation greatly to use the same symbol for an operator and its kernel, for example, $\hat{D}(t,s)$ [or $\hat{B}(s)$] refer to the kernel functions representing the operator $\hat{D}$ (or $\hat{B}$). For operators that map a function space to $\mathbb{R}^n$, such as $\hat{B}$ above, we generally use s (or $\hat{s}$) to denote the variable of the kernel function, and for operators that map $\mathbb{R}^n$ to a function space such as $\hat{C}$ above, we use the variable t (or $\hat{t}$). The kernel representation for the operators $\hat{B}, \hat{C}, \hat{D}$ means that their action is given by

$$\hat{B}u = \int_0^\tau \hat{B}(\hat{s})u(\hat{s})\,d\hat{s} \quad (\hat{C}x)(\hat{t}) = \hat{C}(\hat{t})x, \quad \hat{t} \in [0,\tau]$$

$$(\hat{D}u)(\hat{t}) = \int_0^\tau \hat{D}(\hat{t},\hat{s})u(\hat{s})\,d\hat{s}.$$

Note that the state space of $\hat{G}$ is finite dimensional (the n_x in $\mathbb{R}^{n_x}$ refers to the dimension of the state space of G), while its input and output spaces are infinite dimensional. This fact is significant in that, although lifted systems have infinite-dimensional input and output spaces, they can be realized with a state space of dimension no larger than the dimension of the original continuous-time state space model.

To apply the lifting to the sampled-data problem, consider again the standard problem of Fig. 1, and denote the closed-loop operator by $\mathcal{F}(G, \mathcal{H}_\tau C \mathcal{S}_\tau)$. Since the lifting is an isometry, we have that $\|\mathcal{F}(G,\mathcal{H}_\tau C\mathcal{S}_\tau)\| = \|W_\tau \mathcal{F}(G,\mathcal{H}_\tau C\mathcal{S}_\tau)W_\tau^{-1}\|$, this is shown in Fig. 3(a). In Fig. 3(b), we lump the lifting operators W_τ and W_τ^{-1} and the sample and hold operators and consider a new generalized plant $\tilde{G}$. $\tilde{G}$ is a discrete-time system with one infinite-dimensional input and output (corresponding to $\hat{w}$ and $\hat{z}$) and one finite-dimensional input and output (corresponding to $\tilde{u}$ and $\tilde{y}$). Thus, $\mathcal{F}(\tilde{G}, C) = W_\tau \mathcal{F}(G, \mathcal{H}_\tau C \mathcal{S}_\tau) W_\tau^{-1}$, which means that the closed-loop operator $\mathcal{F}(\tilde{G}, C)$ is in fact the lifting of the closed-loop operator $\mathcal{F}(G, \mathcal{H}_\tau C \mathcal{S}_\tau)$. Since the lifting W_τ is an isometry, we have then characterized the L^∞-induced norm of the hybrid system as the $l^\infty_{L^\infty[0,\tau]}$-induced norm of the time-invariant system $\mathcal{F}(\tilde{G}, C)$. The conclusion is that the problem of minimizing the L^∞ induced norm of the sampled-data system, is equivalent to that of minimizing the induced norm of the infinite dimensional but time-invariant system $\mathcal{F}(\tilde{G}, C)$. The previous discussion together with the characterization of internal stability for hybrid systems in [12] (conditions for nonpathological sampling) yields the following theorem.

Theorem 1: Let G and $\tilde{G}$ be as in Fig. 3, then for any finite dimensional C.

i) $\mathcal{F}(G, \mathcal{H}_\tau C \mathcal{S}_\tau)$ is internally stable if and only if $\mathcal{F}(\tilde{G}, C)$ is.
ii) $\|\mathcal{F}(G, \mathcal{H}_\tau C \mathcal{S}_\tau)\| = \|\mathcal{F}(\tilde{G}, C)\|$.

This reformulation of the sampled-data problem to the problem with $\tilde{G}$ has several advantages, first, the controller has no "structural constraints" on it, in contrast to the previous formulation where the controller is constrained to be a sampled-data controller, i.e., of the form $\mathcal{H}_\tau C \mathcal{S}_\tau$, second, both the controller C and the generalized plant $\tilde{G}$ are shift invariant, thus, the periodicity of the original system is "removed," and third, all parts of the system are operating over the same time set (discrete time). The price paid for these advantages is the infinite dimensionality of the input and output spaces. In this paper, we will show how one can reduce the problem to a finite-dimensional one by "approximating" the input and output spaces by finite-dimensional spaces, thus, reducing the problem to a standard finite-dimensional l^1 problem.

We now present (from [1]) a state space realization for the new generalized plant $\tilde{G}$ which will be useful in studying the problem further. Let the original continuous-time plant G be given by the following realization

$$G = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & 0 & 0 \end{array}\right].$$

It is assumed that the sampler is preceded with a presampling filter which is a strictly causal linear system, this is a realistic assumption since an ideal sampler is not a physical device, a real sampler can be modeled as an integrator with a fast time constant followed by an ideal sampler.

The system shown above represents a generalized plant with the presampling filter absorbed in it, the fact that $D_{21} = D_{22} = 0$ is due to the strict causality of the presampling filter, this also guarantees that the ideal sampler only operates on continuous signals. It can be shown ([1]) that a realization for the generalized plant $\tilde{G}$ (Fig. 3) is given by

$$\tilde{G} = \begin{bmatrix} \tilde{G}_{11} & \tilde{G}_{12} \\ \tilde{G}_{21} & \tilde{G}_{22} \end{bmatrix} = \left[\begin{array}{c|cc} \hat{A} & \hat{B}_1 & \hat{B}_2 \\ \hline \hat{C}_1 & \hat{D}_{11} & \hat{D}_{12} \\ \hat{C}_2 & 0 & 0 \end{array}\right] = \left[\begin{array}{c|cc} e^{A\tau} & e^{A(\tau-s)}B_1 & \Psi(\tau)B_2 \\ \hline C_1e^{At} & C_1e^{A(t-s)}\mathbf{1}(t-s)B_1 + D_{11}\delta(t-s) & C_1\Psi(t)B_2 + D_{12} \\ C_2 & 0 & 0 \end{array}\right]$$

where $\Psi(t) := \int_0^t e^{As}\,ds$. The system G has the following input and output spaces

$$\begin{aligned} \tilde{G}_{11}&: \quad l_{L^\infty[0,\tau]} \to l_{L^\infty[0,\tau]} \\ \tilde{G}_{12}&: \quad l_{R^{n_u}} \to l_{L^\infty[0,\tau]} \\ \tilde{G}_{21}&: \quad l_{L^\infty[0,\tau]} \to l_{R^{n_y}} \\ \tilde{G}_{22}&: \quad l_{R^{n_u}} \to l_{R^{n_y}}. \end{aligned}$$

The main theme of this paper is to approximate the infinite-dimensional input and output spaces $L^\infty[0,\tau]$ by finite-dimensional spaces. Bounds on the approximation of the closed-loop system (i.e., with controller) will be obtained that are characterized only in terms of the operators $\hat{B}_1, \hat{C}_1, \hat{D}_{12}, \hat{D}_{11}$, which in turn are characterized by the original continuous-time plant and independent of the controller.

The interpretation that can be given to the operators $\hat{B}_1, \hat{C}_1, \hat{D}_{12}, \hat{D}_{11}$ is that they characterize the inter-sample behavior of the overall system. In the lifted formulation of the sampled-data problem, the state of the system is the state of the plant $\tilde{G}$ and the state of the controller C, both of which evolve in discrete time. The controller thus has an effect on the state of the system only at the sampling instants, and the inter-sample behavior is governed only by the plant dynamics. This fact is made intuitive by the observation that in between the samples, the system is essentially operating in open loop since there is no feedback (u is constant in between samples).

The lifting of the sampled-data problem makes clear that the inter-sample dynamics are characterized by the operators $\hat{B}_1, \hat{C}_1, \hat{D}_{12}, \hat{D}_{11}$, and thus the issue of approximating these dynamics essentially amounts to approximating the operators, which are independent of the controller. The foregoing ideas are pursued in the next sections.

III. Solution Procedure

Using the lifting we are able to convert the problem of finding a controller to minimize the L^∞ induced norm of the hybrid system (Fig. 1) into the following standard problem with an infinite-dimensional generalized plant $\tilde{G}$:

$$\begin{aligned} \gamma_{\text{opt}} &:= \inf_{C \text{ stabilizing}} \|\mathscr{F}(G, \mathscr{H}_\tau C \mathscr{S}_\tau)\| \\ &= \inf_{C \text{ stabilizing}} \|\mathscr{F}(\tilde{G}, C)\|. \end{aligned} \tag{2}$$

We also note that because of Theorem 1, suboptimal solutions to the above problem will also be suboptimal (with the same norm) for the hybrid system.

The above infinite-dimensional problem is solved by an approximation procedure through solving a standard MIMO l^1 problem. The idea we use is similar to that in [10] and [14] where multirate sampling is used to obtain discrete-time systems that approximate the continuous-time behavior of hybrid systems. This approximation procedure was used in [10] to address the l^1 sampled-data problem. The approximation procedure we use is essentially equivalent to that in [10], however, since we introduce it directly as an approximation to the lifted problem (2), the nature of the approximation is more transparent and we are able to explicitly isolate the parts of the system that need to be approximated independently of the controller. The consequence is that we are able to obtain explicit bounds on the degree of approximation in terms of constants that can be computed *a priori*, and that are dependent only on the plant.

We now describe the approximation procedure. Let $\mathscr{H}_n$ and $\mathscr{S}_n$ be the following operators defined between $L_q^\infty[0,\tau]$ and $l_q^\infty(n)$($l_q^\infty(n)$ is $\mathbb{R}^{n\times q}$ with the maximum norm

$$\mathscr{S}_n: \quad L_q^\infty[0,\tau] \to l_q^\infty(n) \quad (\mathscr{S}_n u)(i) = u\left(\frac{\tau}{n}i\right); \qquad u \in L_q^\infty[0,\tau]$$

$$\mathscr{H}_n: \quad l_q^\infty(n) \to L_q^\infty[0,\tau] \quad (\mathscr{H}_n u)(t) = u\left(\left\lfloor \frac{tn}{\tau} \right\rfloor\right); \qquad \{u(i)\} \in l_q^\infty(n)$$

(strictly speaking, $\mathscr{S}_n$ is not an operator on L_q^∞ but on the subspace of left and right continuous functions, this distinction is irrelevant here since in our setting, assumptions are made to guarantee that $\mathscr{S}_n$ operates only on continuous signals), the above operators can be thought of as "fast" sample and hold operators (see Fig. 5). For simplicity of notation we will suppress the dimension q in the sequel.

Now to approximate the infinite-dimensional problem, we use the approximate closed-loop system $\mathscr{S}_n\mathscr{F}(\tilde{G}, C)\mathscr{H}_n$ (see Fig. 4), and for each n we define

$$\gamma_n := \inf_{C \text{ stabilizing}} \|\mathscr{S}_n\mathscr{F}(\tilde{G}, C)\mathscr{H}_n\|. \tag{3}$$

This new problem now involves the induced norm over

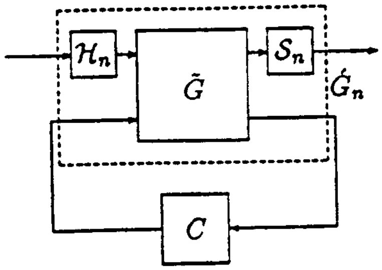


Fig. 4. The system $\acute{G}_n$.

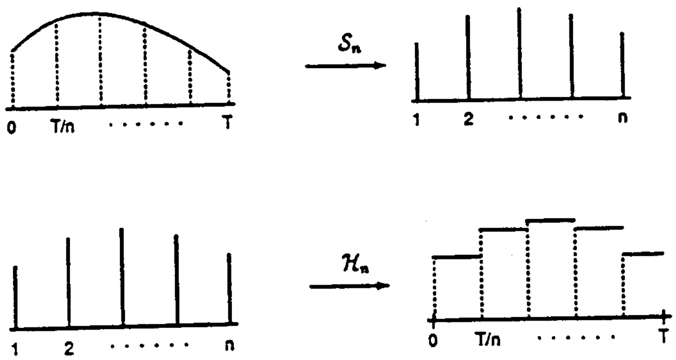


Fig. 5. The operators $\mathcal{S}_n$ and $\mathcal{H}_n$.

$l^\infty_{\Gamma(n)}$, i.e., it is a standard MIMO l^1 problem.

Let us denote the generalized plant associated with $\mathcal{S}_n\mathcal{F}(\bar{G}, C)\mathcal{H}_n$ by $\acute{G}_n$, that is, $\acute{G}_n$ is such that (see Fig. 4)

$$\mathcal{S}_n\mathcal{F}(\bar{G}, C)\mathcal{H}_n = \mathcal{F}(\acute{G}_n, C).$$

A realization for $\acute{G}_n$ is given by,

$$\acute{G}_n = \left[\begin{array}{c|cc} \hat{A} & \hat{B}_1\mathcal{H}_n & \bar{B}_2 \\ \hline \mathcal{S}_n\hat{C}_1 & \mathcal{S}_n\hat{D}_{11}\mathcal{H}_n & \mathcal{S}_n\bar{D}_{12} \\ \bar{C}_2 & 0 & 0 \end{array}\right] =: \left[\begin{array}{c|cc} \hat{A} & \acute{B}_1 & \bar{B}_2 \\ \hline \acute{C}_1 & \acute{D}_{11} & \acute{D}_{12} \\ \bar{C}_2 & 0 & 0 \end{array}\right].$$

The new operators, which are now matrices, are computed to be

$$\acute{C}_1 = \begin{bmatrix} C_1 \\ C_1 e^{A\tau/n} \\ \vdots \\ C_1(e^{A\tau/n})^{n-1} \end{bmatrix}$$

$$\acute{D}_{12} = \begin{bmatrix} D_{12} \\ C_1\Psi(\tau/n)B_2 + D_{12} \\ \vdots \\ C_1\Psi\left(\frac{\tau(n-1)}{n}\right)B_2 + D_{12} \end{bmatrix},$$

$$\acute{D}_{11} = \left\{\left[\begin{array}{c|c} e^{A\tau/n} & \Psi(\tau/n)B_1 \\ \hline C_1 & D_{11} \end{array}\right]\right\}_n$$

$$\acute{B}_1 = \Psi(\tau/n)\left[B_1 \quad e^{A\tau/n}B_1 \quad \cdots \quad (e^{A\tau/n})^{n-1} \quad B_1\right]$$

where $\{\cdot\}_n$ means the first $n \times n$ blocks of the impulse response matrix of the discrete-time system given by the realization in $\{\cdot\}$.

The solution to the original infinite-dimensional problem (and thus to the sampled-data problem) is as follows: n can be chosen large enough such that if the designed controller C_n is almost optimal for the approximate problem (3), then it is almost optimal for the original problem (2). In essence, this approximation scheme "converges," i.e., one can obtain almost optimal controllers by choosing n large enough and solving a MIMO l^1 problem. Exactly what convergence means here is described next.

IV. Design Bounds

In this section we investigate the nature of the approximation of $\|\mathcal{F}(\bar{G}, C)\|$ by $\|\mathcal{F}(\acute{G}_n, C)\|$. In order to show that the synthesis procedure outlined in the previous section yields controllers with performance arbitrarily close to the optimal, one needs to obtain explicit bounds on the degree of approximation of $\|\mathcal{F}(\bar{G}, C)\|$ by $\|\mathcal{F}(\acute{G}_n, C)\|$.

Let us begin with analysis. Note that since $\|\mathcal{F}(\bar{G}, C)\|$ is an infinite-dimensional system, its $l^\infty_{L^\infty[0,\tau]}$-induced norm is not readily computable. A method of computing $\|\mathcal{F}(\bar{G}, C)\|$ comes from the limit

$$\|\mathcal{F}(\bar{G}, C)\| = \lim_{n\to\infty} \|\mathcal{S}_n\mathcal{F}(\bar{G}, C)\mathcal{H}_n\| =: \lim_{n\to\infty} \|\mathcal{F}(\acute{G}_n, C)\| \tag{4}$$

for a fixed C. This formula can be proved using arguments about the approximation of continuous functions by simple functions in L^∞ ([19]), and also follows immediately from the main inequality below. Since $\mathcal{F}(\acute{G}_n, C)$ is a time-invariant MIMO system and $\|\mathcal{F}(\acute{G}_n, C)\|$ is its l^1 norm, it can be computed to any desired accuracy, consequently, by (4) the actual norm, $\|\mathcal{F}(\bar{G}, C)\|$ can be computed to any desired accuracy. However, (4) is by far not sufficient to show the convergence of the synthesis procedure, since given only (4), the rate of convergence may depend on the choice of C.

Our objective is to obtain explicit bounds on $\|\mathcal{F}(\bar{G}, C)\|$ that do not depend on the controller in the following form

Main Inequality: There are constants K_o and K_1 which depend only on G, such that for $n \geq 2n_x$, and τ/n nonpathological

$$\begin{aligned} \|\mathcal{F}(\acute{G}_n, C)\| &\leq \|\mathcal{F}(\bar{G}, C)\| \\ &\leq \frac{K_1}{n} + \left(1 + \frac{K_o}{n}\right)\|\mathcal{F}(\acute{G}_n, C)\|. \end{aligned} \tag{5}$$

Remarks:

a) The significance of the bound (5) is that it is exactly what is needed for synthesis. When one performs an l^1 design on the approximate discretization $\acute{G}_n$, the result is a controller that keeps $\|\mathcal{F}(\acute{G}_n, C)\|$ small, but the objective is to keep the L^∞-induced norm of the hybrid system (or equivalently $\|\mathcal{F}(\bar{G}, C)\|$) small, and the inequality (5) guarantees this. It is thus essential that we bound the hybrid norm from *above* by a function of $\|\mathcal{F}(\acute{G}_n, C)\|$.

b) The above inequality shows that the approximation converges at a rate of $(1/n)$.

The first inequality in (5) is easy to obtain, first note that

$$\|\mathscr{F}(\acute{G}_n, C)\| \le \|\mathscr{F}(\tilde{G}, C)\| \qquad \forall n,$$

since

$$\begin{aligned}\|\mathscr{F}(\acute{G}_n, C)\| &= \|\mathscr{S}_n \mathscr{F}(\tilde{G}, C)\mathscr{H}_n\| \\ &\le \|\mathscr{S}_n\| \|\mathscr{F}(\tilde{G}, C)\| \|\mathscr{H}_n\| \le \|\mathscr{F}(\tilde{G}, C)\|\end{aligned}$$

because $\|\mathscr{H}_n\| \le 1$ on $l^\infty(n)$ and $\|\mathscr{S}_n\| \le 1$ on the subspace of L^∞ for which it is defined.

One way to utilize the main inequality for getting *a priori* guarantees on the hybrid norm in terms of the discrete-time l^1 problem is guided by the following; for a fixed n, if one performs a MIMO l^1 design (as in [9], [17]) on $\acute{G}_n$ and obtains a $\gamma_n + \epsilon$ optimal controller (given by C_n), i.e., $\|\mathscr{F}(\acute{G}_n, C_n)\| \le \gamma_n + \epsilon$, then inequality (5) provides that if C_n is implemented in the hybrid system, then

$$\begin{aligned}\gamma_{\text{opt}} \le \|\mathscr{F}(\tilde{G}, G_n)\| &\le \frac{K_1}{n} + \left(1 + \frac{K_o}{n}\right)\|\mathscr{F}(\acute{G}_n, C)\| \\ &\le \frac{K_1}{n} + \left(1 + \frac{K_o}{n}\right)(\gamma_n + \epsilon) \\ &\le \frac{K_1}{n} + \left(1 + \frac{K_o}{n}\right)(\gamma_{\text{opt}} + \epsilon) \end{aligned} \tag{6}$$

where the last inequality follows from $\gamma_n \le \gamma_{\text{opt}}$, which is a consequence of the first inequality in (5).

The above inequality can be simplified by using an upper bound on γ_{opt}, such a bound can be obtained by finding any stabilizing controller C_o and computing an upper bound on the hybrid norm of $\mathscr{F}(\tilde{G}, C_o)$ (by using the main inequality with a large n). Call that upper bound M. Then by using $\gamma_{\text{opt}} \le M$, inequality (6) can be rewritten as

$$\gamma_{\text{opt}} \le \|\mathscr{F}(\tilde{G}, C_n)\| \le \frac{K_1 + K_o(M + \epsilon)}{n} + \epsilon + \gamma_{\text{opt}}.$$

Thus, in order that C_n guarantees $\|\mathscr{F}(G, \mathscr{H}_\tau C_n \mathscr{S}_\tau)\| \le \gamma_{\text{opt}} + \delta$ for any $\delta > 0$, we choose ϵ and n *a priori* to satisfy

$$\delta \le \frac{K_1 + K_o(M + \epsilon)}{n} + \epsilon.$$

It is worthwhile noting that the problem of minimizing $\|\mathscr{F}(\acute{G}_n, C)\|$ is immediately a standard l^1 problem with time-invariant plant. Also, we note that even though the approximation problem is essentially equivalent to a multirate sampled-data problem, it reflects no structural constraints on the controller. General multirate sampled problems do not share this property (see [7]).

The next section is devoted to the derivation of the main inequality (5). Several interesting issues come up, and we get bounds on the approximation by characterizing the approximation of the infinite-dimensional parts of $\tilde{G}$, namely the operators $\hat{B}_1, \hat{C}_1, \tilde{D}_{12}, \hat{D}_{11}$.

V. Decomposition and Approximation of $\tilde{G}$

It will be very helpful in the derivation of (5) to introduce a decomposition of the infinite-dimensional system $\tilde{G}$ by "extracting" the infinite-dimensional parts of the system. The basic idea is roughly that the behavior of the hybrid system between samples is essentially governed by the infinite-dimensional parts of $\tilde{G}$, namely the operators $\hat{B}_1$, $\hat{C}_1$, $\tilde{D}_{12}$, and $\hat{D}_{11}$. These operators are independent of the controller, and thus it should be possible to approximate the behavior in between the samples independently of the controller by "approximating" the aforementioned operators. To illustrate this point further, we first decompose $\tilde{G}$ as

$$\tilde{G} = \tilde{G}_o + \begin{bmatrix} \hat{D}_{11} & 0 \\ 0 & 0 \end{bmatrix}, \qquad \tilde{G}_o := \left[\begin{array}{c|cc} \hat{A} & \hat{B}_1 & \tilde{B}_2 \\ \hline \hat{C}_1 & 0 & \tilde{D}_{12} \\ \tilde{C}_2 & 0 & 0 \end{array}\right]$$

and we note that $\tilde{G}_o$ can be further decomposed as

$$\tilde{G}_0 = \begin{bmatrix} \begin{bmatrix} \hat{C}_1 & \tilde{D}_{12} \end{bmatrix} & 0 \\ 0 & I \end{bmatrix} \left[\begin{array}{c|cc} \hat{A} & I & \tilde{B}_2 \\ \hline \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} \\ \tilde{C}_2 & 0 & 0 \end{array}\right] \begin{bmatrix} \hat{B}_1 & 0 \\ 0 & I \end{bmatrix}. \tag{7}$$

This decomposition is illustrated in Fig. 6. The closed-loop mapping $\mathscr{F}(\tilde{G}, C)$ is correspondingly decomposed as

$$\begin{aligned}\mathscr{F}(\tilde{G}, C) &= \hat{D}_{11} + \mathscr{F}(\tilde{G}_o, C) \\ &= \hat{D}_{11} + \begin{bmatrix} \hat{C}_1 & \tilde{D}_{12} \end{bmatrix} \mathscr{F}(\tilde{G}_{oo}, C)\hat{B}_1. \end{aligned} \tag{8}$$

We will use the notation $\hat{\mathscr{C}} := [\hat{C}_1 \;\; \tilde{D}_{12}]$, and call $\hat{\mathscr{C}}$ the output operator and $\hat{B}_1$ the input operator.

With this decomposition, $\tilde{G}_{oo}$ is finite dimensional, and $\hat{\mathscr{C}}, \hat{B}_1$ are finite rank operators

$$\hat{\mathscr{C}}: \mathbb{R}^{n_x + n_u} \to L^\infty[0, \tau], \qquad \hat{B}_1: L^\infty[0, \tau] \to \mathbb{R}^{n_x}.$$

As (8) shows, only a finite-dimensional part of the system [i.e., $\mathscr{F}(G(\tilde{G}_{oo}, C)$] is dependent on the controller, while the infinite-dimensional parts are independent of C. Roughly speaking, the controller (being discrete time) only effects the hybrid system at the sampling instants, while in between the samples, the systems evolution is governed by the operators $\hat{D}_{11}, \hat{\mathscr{C}}, \hat{B}_1$, which are in turn dependent only on the dynamics of the original generalized plant G.

The remainder of this section and the appendixes are devoted to deriving the main inequality, and can be skipped without loss of continuity.

We now consider the issue of "approximating" the infinite-dimensional plant $\tilde{G}$ by a finite-dimensional plant $\acute{G}_n$. First we note that the two norms to be compared are

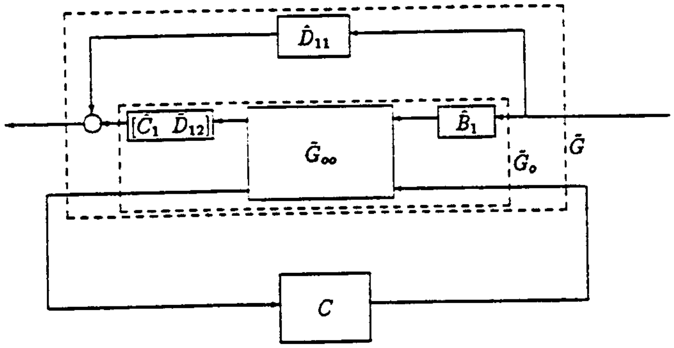


Fig. 6. Decomposition of $\tilde{G}$.

of $\mathcal{F}(\tilde{G}, C)$, which has $L^\infty[0, \tau]$ as an input–output space, and of $\mathcal{S}_n\mathcal{F}(\tilde{G}, C)\mathcal{H}_n$, which has $l^\infty(n)$ as an input–output space. Therefore, it is not strictly true that $\mathcal{S}_n\mathcal{F}(\tilde{G}, C)\mathcal{H}_n$ approximates $\mathcal{F}(\tilde{G}, C)$ since comparisons like $\|\mathcal{F}(\tilde{G}, C) - \mathcal{S}_n\mathcal{F}(\tilde{G}, C)\mathcal{H}_n\| \leq \epsilon$ do not make sense. We will replace $\mathcal{S}_n\mathcal{F}(\tilde{G}, C)\mathcal{H}_n$ by another system which has the same norm, but truly approximates $\mathcal{F}(\tilde{G}, C)$.

Define the following operator (the normalized integration operator) $\mathcal{J}_n$: $L^\infty[0, \tau] \to l^\infty(n)$ by

$$(\mathcal{J}_n(u))(i) := \frac{n}{\tau}\int_{i\tau/n}^{(i+1)\tau/n} u(t)\, dt.$$

The following properties of $\mathcal{J}_n$ can be easily checked: $\mathcal{J}_n$ is a linear operator, $\|\mathcal{J}_n\| = 1$, and $\mathcal{J}_n$ is a left inverse to $\mathcal{H}_n$, i.e. $\mathcal{J}_n\mathcal{H}_n$ = identity. If $\mathcal{J}_n$ is regarded as an operator on $L^1[0, \tau]$, i.e., $\mathcal{J}_n$: $L^1[0, \tau] \to l^1(n)$, then it is easily shown that $\mathcal{H}_n$ is the adjoint of $(\tau/n)\mathcal{J}_n$, that is $((\tau/n)\mathcal{J}_n)^* = \mathcal{H}_n$. Similarly, if $\mathcal{H}_n$ is regarded as an operator on $l^1(n)$, i.e., $\mathcal{H}_n$: $l^1(n) \to L^1[0, \tau]$, then $\mathcal{H}_n^* = (\tau/n)\mathcal{J}_n$, which also implies that $(\mathcal{H}_n\mathcal{J}_n)^* = \mathcal{H}_n\mathcal{J}_n$.

Let us denote by $T := \mathcal{F}(\tilde{G}, C)$, and by $\acute{T}_n := \mathcal{F}(\acute{G}_n, C)$. As already mentioned, T and $\acute{T}_n$ cannot be compared directly since they do not have the same input and output space. The operator $\mathcal{J}_n$ will allow us to form a system $\overline{T}_n$ with norm equal to that of $\acute{T}_n$, but with the same input and output spaces as T.

Lemma 2: Define the system $\overline{T}_n := (\mathcal{H}_n\mathcal{S}_n)T(\mathcal{H}_n\mathcal{J}_n)$, then

$$\|\overline{T}_n\| = \|\acute{T}_n\|.$$

Proof: It is true that $\|\mathcal{S}_n T\mathcal{H}_n\mathcal{J}_n\| = \|\mathcal{S}_n T\mathcal{H}_n\|$ since

$$\|\mathcal{S}_n T\mathcal{H}_n\mathcal{J}_n\| \leq \|\mathcal{S}_n T\mathcal{H}_n\|\,\|\mathcal{J}_n\| \leq \|\mathcal{S}_n T\mathcal{H}_n\|,$$

and

$$\|\mathcal{S}_n T\mathcal{H}_n\| \leq \|\mathcal{S}_n T\mathcal{H}_n\mathcal{J}_n\mathcal{H}_n\| \leq \|\mathcal{S}_n T\mathcal{H}_n\mathcal{J}_n\|.$$

Also, since $\mathcal{H}_n$: $l^\infty(n) \to L^\infty[0, \tau]$ is an isometry, we conclude that

$$\|\overline{T}_n\| := \|\mathcal{H}_n\mathcal{S}_n T\mathcal{H}_n\mathcal{J}_n\| = \|\mathcal{S}_n T\mathcal{H}_n\mathcal{J}_n\| = \|\mathcal{S}_n T\mathcal{H}_n\| =: \|\acute{T}_n\|.$$

■

Remark: The above lemma is of general interest since it provides a systematic way of addressing the question of how a discretized system $\mathcal{S}_\tau H\mathcal{H}_\tau$ "approximates" the original system H, by comparing the systems H and $\overline{H} := (\mathcal{H}_\tau\mathcal{S}_\tau)H(\mathcal{H}_\tau\mathcal{J}_\tau)$. This comparison is typically easier since H and $\overline{H}$ are both continuous-time systems with the same input and output spaces.

Let $\overline{G}_n$ be the generalized plant corresponding to the closed-loop operator $\overline{T}_n$, i.e., $\overline{T}_n = \mathcal{F}(\overline{G}_n, C)$. $\overline{G}_n$ is defined by

$$\overline{G}_n = \begin{bmatrix} \mathcal{H}_n\mathcal{S}_n & 0 \\ 0 & I \end{bmatrix} \tilde{G} \begin{bmatrix} \mathcal{H}_n\mathcal{J}_n & 0 \\ 0 & I \end{bmatrix}.$$

The consequence of Lemma 2 is that one only needs to show inequality (5) with $\mathcal{F}(\overline{G}_n, C)$ instead of $\mathcal{F}(\acute{G}_n, C)$. As already mentioned, the advantage is that $\mathcal{F}(\overline{G}_n, C)$ has the same input and output spaces as $\mathcal{F}(\tilde{G}, C)$, namely $L^\infty[0, \tau]$.

Next, we will show that $\mathcal{F}(\overline{G}_n, C)$ actually approximates $\mathcal{F}(\tilde{G}, C)$, and this will yield the main inequality (5).

Approximation of $\tilde{G}$: The approximation of $\tilde{G}$ will be done in two parts corresponding to the decomposition $\mathcal{F}(\tilde{G}, C) = \hat{D}_{11} + \mathcal{F}(\tilde{G}_o, C) = \hat{D}_{11} + \hat{\mathscr{C}}\mathcal{F}(\tilde{G}_{oo}, C)\hat{B}_1$. It will be useful in this section to use a short hand notation for (see Fig. 7)

$$T_o := \hat{\mathscr{C}}\mathcal{F}\big(\tilde{G}_{oo}, C\big)\hat{B}_1 \qquad T_{oo} := \mathcal{F}\big(\tilde{G}_{oo}, C\big) \tag{9}$$

$$\overline{T}_{on} := (\mathcal{H}_n\mathcal{S}_n)T_o(\mathcal{H}_n\mathcal{J}_n) \qquad \overline{D}_n := (\mathcal{H}_n\mathcal{S}_n)\hat{D}_{11}(\mathcal{H}_n\mathcal{J}_n) \tag{10}$$

and corresponding to the decomposition $T = \hat{D}_{11} + T_{on}$, we have

$$\overline{T}_n = (\mathcal{H}_n\mathcal{S}_n)\big(\hat{D}_{11} + T_o\big)(\mathcal{H}_n\mathcal{J}_n) = \overline{D}_n + \overline{T}_{on}.$$

We will first show that $\overline{T}_{on}$ approximates T_o, then we show that $\overline{D}_n$ approximates $\hat{D}_{11}$.

Proposition 3: Let $n \geq 2n_x$, such that τ/n is not a pathological sampling period, there exists a constant K_o which depends only on G, such that

$$\|T_o - \overline{T}_{on}\| \leq \frac{K_o}{n}\|\overline{T}_{on}\|.$$

Remark: It is important that the above bound is in terms of $\|\overline{T}_{on}\|$ which corresponds to part of $\mathcal{F}(\acute{G}_n, C)$. The reason being that in the main inequality, we must bound the norm of the hybrid system from *above* by the norm of the discretized system $\mathcal{F}(\acute{G}_n, C)$. In fact, it is much easier to produce an inequality as above but with $\|T_o\|$ on the right-hand side, but this would not be useful for bounding the norm of the hybrid system.

Proof: The proof makes use of the decomposition of $T_o = \hat{\mathscr{C}}T_{oo}\hat{B}_1$, and of its approximation $\overline{T}_{on} = (\mathcal{H}_n\mathcal{S}_n)\hat{\mathscr{C}}T_{oo}\hat{B}_1(\mathcal{H}_n\mathcal{J}_n)$. The basic idea of the proof (on the output side) is that $(\mathcal{H}_n\mathcal{S}_n)$ operates on functions in $\mathscr{R}_{(\hat{\mathscr{C}})} \subset L^\infty[0, \tau]$, and functions in $\mathscr{R}_{(\hat{\mathscr{C}})}$ are continuous and there are bounds on their rate of change (depending on the dynamics of the plant), so on $\mathscr{R}_{(\hat{\mathscr{C}})}$ the operator $(\mathcal{H}_n\mathcal{S}_n)$ approximates the identity, and it also has a left inverse which approximates the identity as $n \to \infty$.

We now approximate from the output side. Lemma 4 below states that $(\mathcal{H}_n\mathcal{S}_n)$ has a left inverse on $\mathscr{R}_{(\hat{\mathscr{C}})}$, i.e., there exists $(\mathcal{H}_n\mathcal{S}_n)^{-L}$: $\mathscr{R}_{(\mathcal{H}_n\mathcal{S}_n\hat{\mathscr{C}})} \to \mathscr{R}_{(\hat{\mathscr{C}})} \subset L^\infty[0, \tau]$

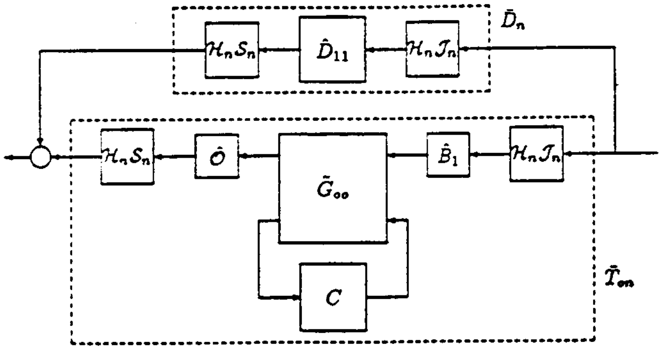


Fig. 7. Decomposition of the approximate system $\mathcal{S}(\overline{G}_n, C)$.

such that $(\mathcal{H}_n\mathcal{S}_n)^{-L}(\mathcal{H}_n\mathcal{S}_n) = $ identity on $\mathcal{R}_{(\hat{\mathcal{O}})}$. We now establish

$$\begin{aligned}\|(\mathcal{H}_n\mathcal{S}_n)T_o - T_o\| &= \|(\mathcal{H}_n\mathcal{S}_n)\hat{\mathcal{O}}T_{oo}\hat{B}_1 - \big((\mathcal{H}_n\mathcal{S}_n)^{-L}(\mathcal{H}_n\mathcal{S}_n)\hat{\mathcal{O}}T_{oo}\hat{B}_1\big)\| \\ &\le \|\big(I - (\mathcal{H}_n\mathcal{S}_n)^{-L}\big)|_{\mathcal{R}(\mathcal{H}_n\mathcal{S}_n\hat{\mathcal{O}})}\| \, \|(\mathcal{H}_n\mathcal{S}_n)T_o\|\end{aligned}$$

where the operator I is the identity, or the embedding I: $\mathcal{R}_{(\mathcal{H}_n\mathcal{S}_n\hat{\mathcal{O}})} \to L^\infty[0, \tau]$. Also from Lemma 4, we have that $\|(I - (\mathcal{H}_n\mathcal{S}_n)^{-L})|_{\mathcal{R}(\mathcal{H}_n\mathcal{S}_n\hat{\mathcal{O}})}\| \le (K_{\hat{\mathcal{O}}}/n)$, this implies

$$\|(\mathcal{H}_n\mathcal{S}_n)T_o - T_o\| \le \frac{K_{\hat{\mathcal{O}}}}{n}\|(\mathcal{H}_n\mathcal{S}_n)T_o\|. \tag{11}$$

Now, to approximate on the input side, we need to take preadjoints (see Appendix B):

$$\begin{aligned}\|(\mathcal{H}_n\mathcal{S}_n)T_o - (\mathcal{H}_n\mathcal{S}_n)T_o(\mathcal{H}_n\mathcal{T}_n)\| &= \|(\mathcal{H}_n\mathcal{S}_n)\hat{\mathcal{O}}T_{oo}\hat{B}_1 - (\mathcal{H}_n\mathcal{S}_n)\hat{\mathcal{O}}T_{oo}\hat{B}_1(\mathcal{H}_n\mathcal{T}_n)\| \\ &= \|(\mathcal{H}_n\mathcal{S}_n)\hat{\mathcal{O}}T_{oo}\big(\hat{B}_1 - \hat{B}_1(\mathcal{H}_n\mathcal{T}_n)\big)\| \\ &= \|\big({}^*\hat{B}_1 - {}^*\mathcal{T}_n^*\mathcal{H}_n^*{}^*\hat{B}_1\big)^*\big((\mathcal{H}_n\mathcal{S}_n)\hat{\mathcal{O}}T_{oo}\big)\| \\ &= \|\big({}^*\hat{B}_1 - (\mathcal{H}_n\mathcal{T}_n)^*\hat{B}_1\big)^*\big((\mathcal{H}_n\mathcal{S}_n)\hat{\mathcal{O}}T_{oo}\big)\|.\end{aligned}$$

From Lemma 4 below, $(\mathcal{H}_n\mathcal{T}_n)$ has a left inverse when restricted to $\mathcal{R}_{({}^*\hat{B}_1)}$, i.e., $(\mathcal{H}_n\mathcal{T}_n)^{-L}$ is such that $(\mathcal{H}_n\mathcal{T}_n)^{-L}(\mathcal{H}_n\mathcal{T}_n) = $ identity on $\mathcal{R}_{({}^*\hat{B}_1)} \subset L^1[0, \tau]$, therefore

$$\begin{aligned}\|(\mathcal{H}_n\mathcal{S}_n)T_o - (\mathcal{H}_n\mathcal{S}_n)T_o(\mathcal{H}_n\mathcal{T}_n)\| &= \|\big((\mathcal{H}_n\mathcal{T}_n)^{-L}(\mathcal{H}_n\mathcal{T}_n)^*\hat{B}_1 \\ &\quad - (\mathcal{H}_n\mathcal{T}_n)^*\hat{B}_1\big)^*((\mathcal{H}_n\mathcal{S}_n)\mathcal{O}T_{oo})\| \\ &= \|\big((\mathcal{H}_n\mathcal{T}_n)^{-L} - I\big)|_{\mathcal{R}(\mathcal{H}_n\mathcal{T}_n{}^*\hat{B}_1)}{}^* \\ &\quad \cdot\big((\mathcal{H}_n\mathcal{S}_n)\hat{\mathcal{O}}T_{oo}\hat{B}_1(\mathcal{H}_n\mathcal{T}_n)\big)\| \\ &\le \|\big((\mathcal{H}_n\mathcal{T}_n)^{-L} - I\big)|_{\mathcal{R}(\mathcal{H}_n\mathcal{T}_n{}^*\hat{B}_1)}\| \, \|(\mathcal{H}_n\mathcal{S}_n)T_o(\mathcal{H}_n\mathcal{T}_n)\| \\ &\le \frac{K_{\hat{B}}}{n}\|\overline{T}_{on}\|\end{aligned} \tag{12}$$

where the last step is again from Lemma 4.

Combining inequalities (11) and (12), we get

$$\begin{aligned}\|T_o - \overline{T}_{on}\| &= \|T_o - (\mathcal{H}_n\mathcal{S}_n)T_o(\mathcal{H}_n\mathcal{T}_n)\| \\ &= \|T_o - (\mathcal{H}_n\mathcal{S}_n)T_o + (\mathcal{H}_n\mathcal{S}_n)T_o \\ &\quad - (\mathcal{H}_n\mathcal{S}_n)T_o(\mathcal{H}_n\mathcal{T}_n)\| \\ &\le \|T_o - (\mathcal{H}_n\mathcal{S}_n)T_o\| + \|(\mathcal{H}_n\mathcal{S}_n)T_o \\ &\quad - (\mathcal{H}_n\mathcal{S}_n)T_o(\mathcal{H}_n\mathcal{T}_n)\| \\ &\le \frac{K_{\hat{\mathcal{O}}}}{n}\|(\mathcal{H}_n\mathcal{S}_n)T_o\| + \frac{K_{\hat{B}}}{n}\|\overline{T}_{on}\|\end{aligned}$$

but (12) also implies that $\|(\mathcal{H}_n\mathcal{S}_n)T_o\| \le (1 + (K_{\hat{B}}/n))\|\overline{T}_{on}\|$, therefore

$$\|T_o - \overline{T}_{on}\| \le \left[\frac{K_{\hat{\mathcal{O}}}}{n}\left(1 + \frac{K_{\hat{B}}}{n}\right) + \frac{K_{\hat{B}}}{n}\right]\|\overline{T}_{on}\| \le \frac{K_o}{n}\|\overline{T}_{on}\|,$$

where $K_o := K_{\hat{\mathcal{O}}} + K_{\hat{\mathcal{O}}}K_{\hat{B}} + K_{\hat{B}}$. ∎

Lemma 4 below captures the idea that $(\mathcal{H}_n\mathcal{S}_n)\hat{\mathcal{O}}$ approximates $\hat{\mathcal{O}}$, because the sampling operator $\mathcal{S}_n$ samples only elements in $\mathcal{R}(\hat{\mathcal{O}})$, and since there is a bound on the variation of functions in $\mathcal{R}(\hat{\mathcal{O}})$, one can get a bound on how well $(\mathcal{H}_n\mathcal{S}_n)$ approximates elements in $\mathcal{R}(\hat{\mathcal{O}})$. Similar arguments are made about $(\mathcal{H}_n\mathcal{T}_n)^*\hat{B}_1$. This lemma is the key to obtaining approximations that are independent of the controllers, since the behavior of the signals in the input and output spaces is governed by $\hat{\mathcal{O}}$ and $\hat{B}_1$, the nature of the approximation depends on these two operators and not C. The rate of convergence of the approximations is determined by the constants $K_{\hat{B}}, K_{\hat{\mathcal{O}}}$, which are completely determined by the operators $\hat{B}$ and $\hat{\mathcal{O}}$, respectively, which in turn, are completely determined by the original plant.

Lemma 4: Assume $n \ge 2n_x$, and τ/n is not a pathological sampling period, then

a) $\exists$ an operator $(\mathcal{H}_n\mathcal{T}_n)^{-L}$: $\mathcal{R}_{(\mathcal{H}_n\mathcal{T}_n{}^*\hat{B}_1)} \to L^1[0, \tau]$ such that $(\mathcal{H}_n\mathcal{T}_n)^{-L}(\mathcal{H}_n\mathcal{T}_n)|_{\mathcal{R}({}^*\hat{B}_1)} = $ identity,

$$\begin{array}{ccccc} L^1[0,\tau] & & L^1[0,\tau] & & L^1[0,\tau] \\ \cup & (\mathcal{H}_n\mathcal{T}_n)^{-L} & \cup & (\mathcal{H}_n\mathcal{T}_n) & \cup \\ \mathcal{R}_{({}^*\hat{B}_1)} & \leftarrow & \mathcal{R}_{(\mathcal{H}_n\mathcal{T}_n{}^*\hat{B}_1)} & \leftarrow & \mathcal{R}_{({}^*\hat{B}_1)} \end{array}$$

and a constant $K_{\hat{B}}$, such that

$$\|\big(I - (\mathcal{H}_n\mathcal{T}_n)^{-L}\big)|_{\mathcal{R}(\mathcal{H}_n\mathcal{T}_n{}^*\hat{B}_1)}\| \le \frac{K_{\hat{B}}}{n}.$$

b) $\exists$ an operator $(\mathcal{H}_n\mathcal{S}_n)^{-L}$: $\mathcal{R}_{(\mathcal{H}_n\mathcal{S}_n\hat{\mathcal{O}})} \to L^\infty[0, \tau]$ such that $(\mathcal{H}_n\mathcal{S}_n)^{-L}(\mathcal{H}_n\mathcal{S}_n)|_{\mathcal{R}(\hat{\mathcal{O}})} = $ identity,

$$\begin{array}{ccccc} L^\infty[0,\tau] & & L^\infty[0,\tau] & & L^\infty[0,\tau] \\ \cup & (\mathcal{H}_n\mathcal{S}_n)^{-L} & \cup & (\mathcal{H}_n\mathcal{S}_n) & \cup \\ \mathcal{R}_{(\hat{\mathcal{O}})} & \leftarrow & \mathcal{R}_{(\mathcal{H}_n\mathcal{S}_n\hat{\mathcal{O}})} & \leftarrow & \mathcal{R}_{(\hat{\mathcal{O}})} \end{array}$$

and a constant $K_{\hat{\mathcal{O}}}$ such that

$$\|\big(I - (\mathcal{H}_n\mathcal{S}_n)^{-L}\big)|_{\mathcal{R}(\mathcal{H}_n\mathcal{S}_n\hat{\mathcal{O}})}\| \le \frac{K_{\hat{\mathcal{O}}}}{n}.$$

The proofs of this lemma and the next one are quite technical and involved, and thus are relegated to the appendix.

The next lemma takes care of approximating the direct feed-through operator $\hat{D}_{11}$, which is approximated by the direct feed-through operator $\overline{D}_n$ of $\overline{G}_n$.

Lemma 5: There is a constant $K_{\hat{D}}$ such that

$$\|\hat{D}_{11} - \overline{D}_n\| \le \frac{K_{\hat{D}}}{n}.$$

Combining Proposition 3 and Lemma 5, we get that $\overline{T}_n$ approximates T by

$$\|T - \overline{T}_n\| \le \frac{K_o}{n}\|\overline{T}_{on}\| + \frac{K_{\hat{D}}}{n}. \tag{13}$$

To get a bound with $\|\overline{T}_n\|$ on the right, note that $\overline{T}_n = \overline{T}_{on} + \overline{D}_n$, which implies by the triangle inequality that $\|\overline{T}_{on}\| - \|\overline{D}_n\| \le \|\overline{T}_n\|$, and

$$\|\overline{T}_{on}\| \le \|\overline{D}_n\| + \|\overline{T}_n\| \le \|\hat{D}_{11}\| + \|\overline{T}_n\|.$$

Since $\|\hat{D}_{11}\|$ is a constant, combining with (13) yields

$$\|T - \overline{T}_n\| \le \frac{K_o}{n}\|\hat{D}_{11}\| + \frac{K_o}{n}\|\overline{T}_n\| + \frac{K_{\hat{D}}}{n}.$$

Finally, since $\|T\| - \|\overline{T}_n\| \le \|T - \overline{T}_n\|$, we get

$$\|T\| \le \frac{K_o\|\hat{D}_{11}\| + K_{\hat{D}}}{n} + \left(1 + \frac{K_o}{n}\right)\|\overline{T}_n\|$$
$$=: \frac{K_1}{n} + \left(1 + \frac{K_o}{n}\right)\|\overline{T}_n\|$$

and thus we have arrived at the main inequality (5).

VI. Geometrical Interpretations

In the previous section we gave an approximation procedure to obtain approximately optimal controllers. The procedure is based on forming an "approximate" finite-dimensional system to an infinite-dimensional one. A question may be asked as to whether the infinite-dimensional problem may be *exactly* reducible to a finite-dimensional l^1 problem. For example, in [1], the $\mathscr{H}^\infty$ sampled-data problem was treated by the lifting technique, and an exact reduction of the resulting infinite-dimensional problem to a finite-dimensional one is possible. This motivates the question as to whether a similar exact reduction is possible in the l^1 problem.

In this section, we will not give a definite answer to this question, but it is our purpose to illustrate some of the underlying geometry in the reduction, and to suggest that the l^1 sampled-data problem may not be exactly reducible to a finite-dimensional l^1 problem. We will give a geometric reasoning which shows that the fundamental difference between the reduction of the $\mathscr{H}^\infty$ and the l^1 sampled-data problems has to do with the difference between the geometry of finite-dimensional Hilbert and Banach spaces.

Let us go back to the formulation of the problem involving the infinite-dimensional generalized plant $\tilde{G}$, and consider the decomposition of $\tilde{G}$ in feedback with the controller C (Fig. 6).

To facilitate the geometric arguments we are about to make, we assume that the operator $\hat{D}_{11} = 0$. Note that this assumption is valid only when $G_{11} = 0$, and this is an unrealistic assumption for most interesting control problems, but the assumption is made for the purpose of illustration. With the assumption $\hat{D}_{11} = 0$, the decomposed system in feedback with C is shown in Fig. 8, where $\hat{\mathscr{O}} := [\hat{C}_1 \;\; \hat{D}_{12}]$.

We first look at possible decompositions of the output space $L^\infty[0,\tau]$. From Fig. 8, it is clear that

$$\mathscr{F}(\tilde{G}, C) = \hat{\mathscr{O}}\mathscr{F}(\tilde{G}_{oo}, C)\hat{B}_1$$

which means that the output signal $\hat{z}$ takes values in $\mathscr{R}(\hat{\mathscr{O}}) \subset L^\infty[0,\tau]$ (at each point in time). Since $\hat{\mathscr{O}}$: $\mathbb{R}^{n_x+n_u} \to L^\infty[0,\tau]$, then $\mathscr{R}(\hat{\mathscr{O}})$ is a finite-dimensional subspace of $L^\infty[0,\tau]$, and there exists a projection on it $\Pi_{\mathscr{R}(\hat{\mathscr{O}})}$: $L^\infty[0,\tau] \to \mathscr{R}(\hat{\mathscr{O}})$ [20]. By the definition of a projection, we have that for any $x \in \mathbb{R}^{n_x+n_u}$, $\|\hat{\mathscr{O}}x\|_{L^\infty[0,\tau]} = \|\Pi_{\mathscr{R}(\hat{\mathscr{O}})}\hat{\mathscr{O}}x\|_{\mathscr{R}(\hat{\mathscr{O}})}$, therefore

$$\left\|\Pi_{\mathscr{R}(\hat{\mathscr{O}})}\hat{\mathscr{O}}\mathscr{F}(\tilde{G}_{oo}, C)\hat{B}_1\right\| = \left\|\hat{\mathscr{O}}\mathscr{F}(\tilde{G}_{oo}, C)\hat{B}_1\right\| = \left\|\mathscr{F}(\tilde{G}, C)\right\|.$$

Note that $\Pi_{\mathscr{R}(\hat{\mathscr{O}})}\hat{\mathscr{O}}\mathscr{F}(\tilde{G}_{oo}, C)\hat{B}_1$ is a system with a finite-dimensional output space, namely $\mathscr{R}(\hat{\mathscr{O}})$, and the norm on $\mathscr{R}(\hat{\mathscr{O}})$ is the norm it inherits as a subspace of $L^\infty[0,\tau]$.

A similar reduction is possible with the input space, for this, we need to look at the preadjoint operators. Since for any Banach space operator A, $\|A\| = \|A^*\|$, we have that

$$\left\|\Pi_{\mathscr{R}(\hat{\mathscr{O}})}\hat{\mathscr{O}}\mathscr{F}(\tilde{G}_{oo}, C)\hat{B}_1\right\| = \left\|{}^*\hat{B}_1\,{}^*\mathscr{F}(\tilde{G}_{oo}, C)\,{}^*\hat{\mathscr{O}}\,{}^*\Pi_{\mathscr{R}(\hat{\mathscr{O}})}\right\|$$

and as before, we can project on $\mathscr{R}({}^*\hat{B}_1) \subset \mathscr{L}^1[0,\tau]$ without changing the induced norm

$$\left\|{}^*\hat{B}_1{}^*\mathscr{F}(\tilde{G}_{oo}, C){}^*\hat{\mathscr{O}}{}^*\Pi_{\mathscr{R}(\hat{\mathscr{O}})}\right\|$$
$$= \left\|\Pi_{\mathscr{R}({}^*\hat{B}_1)}{}^*\hat{B}_1{}^*\mathscr{F}(\tilde{G}_{oo}, C){}^*\hat{\mathscr{O}}{}^*\Pi_{\mathscr{R}(\hat{\mathscr{O}})}\right\|$$
$$= \left\|\Pi_{\mathscr{R}(\hat{\mathscr{O}})}\hat{\mathscr{O}}\mathscr{F}(\tilde{G}_{oo}, C)\hat{B}_1\Pi^*_{\mathscr{R}({}^*\hat{B}_1)}\right\|$$

where the last equality follows by taking the adjoints. Also, note that since $\Pi_{\mathscr{R}({}^*\hat{B}_1)}$: $L^1[0,\tau] \to \mathscr{R}({}^*\hat{B}_1)$ then $\Pi^*_{\mathscr{R}({}^*\hat{B}_1)}$: $(\mathscr{R}({}^*\hat{B}_1))^* \to L^\infty[0,\tau]$, where $(\mathscr{R}({}^*\hat{B}_1))^*$ is the dual space of $\mathscr{R}({}^*\hat{B}_1)$, and it is finite-dimensional since $\mathscr{R}({}^*\hat{B}_1)$ is.

Combining the reduction on both the input and the output spaces, we have

$$\left\|\hat{\mathscr{O}}\mathscr{F}(\tilde{G}_{oo}, C)\hat{B}_1\right\| = \left\|\Pi_{\mathscr{R}(\hat{\mathscr{O}})}\hat{\mathscr{O}}\mathscr{F}(\tilde{G}_{oo}, C)\hat{B}_1\Pi^*_{\mathscr{R}({}^*\hat{B}_1)}\right\|$$
$$=: \left\|\mathscr{F}(\overline{G}, C)\right\|, \tag{14}$$

where $\overline{G}$ is defined by

$$\overline{G} := \begin{bmatrix} \Pi_{\mathscr{R}(\hat{\mathscr{O}})} & 0 \\ 0 & I \end{bmatrix} \tilde{G} \begin{bmatrix} \Pi^*_{\mathscr{R}({}^*\hat{B}_1)} & 0 \\ 0 & I \end{bmatrix}.$$

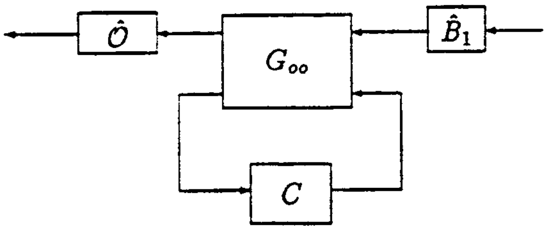


Fig. 8. Decomposition of $\bar{G}$ with $\hat{D}_{11} = 0$.

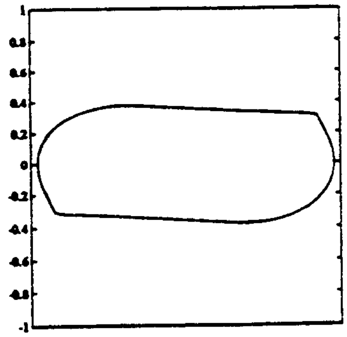


Fig. 9. The unit ball of $\mathscr{R}(\hat{C}_1)$.

Equation (14) shows that the original problem is reducible to the standard problem with the generalized plant $\overline{G}$. Since $\overline{G}$ has finite-dimensional input and output spaces (since $\mathscr{R}(\hat{\mathcal{O}})$ and $(\mathscr{R}(^*\hat{B}_1))^*$ are finite dimensional), we have arrived at an equivalent finite-dimensional problem. This problem is not necessarily a standard finite-dimensional l^1 problem, it is only so if the input and output spaces ($\mathscr{R}(\hat{\mathcal{O}})$ and $(\mathscr{R}(^*\hat{B}_1))^*$) are linearly isometrically isomorphic to an $l^\infty(n)$ space for some n.

Remark: In the $\mathscr{H}^\infty$ sampled-data problem, the situation is much simpler. In that case, $\mathscr{R}(\hat{\mathcal{O}})$ and $(\mathscr{R}(^*\hat{B}_1))^*$ as subspaces of $L^2[0,\tau]$, are immediately linearly isometric to Euclidean spaces (that is $l^2(n)$), since every finite-dimensional Hilbert space is linearly isometric to a Euclidean space of equal dimension.

Thus, the question arises as to what the spaces $\mathscr{R}(\hat{\mathcal{O}})$ and $(\mathscr{R}(^*\hat{B}_1))^*$ look like, and to whether they are isometric to $l^\infty(n)$? If the answer is affirmative, we can use this identification with $l^\infty(n)$ and obtain a generalized plant which has an $l^\infty(n)$ for each of its input and output spaces, and the problem then becomes a standard l^1 problem. However, the answer is negative. This can be seen by a simple example, where we plot the unit ball of the space $\mathscr{R}(\hat{\mathcal{O}})$ and show that there is no linear transformation that can transform it to a unit ball of an $l^\infty(n)$ space.

The example we consider is as follows: first recall that the operator $\hat{\mathcal{O}}$ is given by the following kernel function

$$\hat{\mathcal{O}}(t) := \left[\hat{C}_1(t) \quad \hat{D}_{12}(t)\right] = \left[C_1 e^{At} \quad C_1\left(\int_0^t e^{As}\,ds\right)B_2\right].$$

We will consider the subspace $\mathscr{R}(\hat{C}_1) \subset \mathscr{R}(\hat{\mathcal{O}})$ and show that it cannot be a subspace of any $l^\infty(n)$. Recall that the norm on the space $\mathscr{R}(\hat{C}_1)$ is the norm inherited as a subspace of $L^\infty[0,\tau]$. The unit ball in $\mathscr{R}(\hat{C}_1)$ can be plotted by choosing a basis, and then computing the $L^\infty[0,\tau]$ norm for combinations of the basis elements. The particular example we pick is

$$A = \begin{bmatrix} 0 & -3 \\ 1 & 1 \end{bmatrix}; \qquad C = [1 \quad 1/2],$$

with $\tau = 1$. For this example $\mathscr{R}(\hat{C}_1)$ has dimension two, and a basis for it is given by

$$x_1(t) := C_1 e^{At}\begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad x_2(t) := C_1 e^{At}\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

To plot the unit ball in $\mathscr{R}(\hat{C}_1)$, we represent any $x \in \mathscr{R}(\hat{C}_1)$ by $x = \alpha_1 x_1 + \alpha_2 x_2$. The ball in Fig. 9 represents $\|x\| = 1$, and the axes are α_1 and α_2. The unit ball in an $l^\infty(n)$ space is an n-cube, and the unit ball of any 2-dimensional subspace of $l^\infty(n)$ is a 2-dimensional "slice" through an n-cube, and it is clear that the boundary of this 2-dimensional cube must be made up of straight lines, i.e., it must be a polygon. Now, for $\mathscr{R}(\hat{C}_1)$ to be linearly isometric to a subspace of $l^\infty(n)$, a necessary condition is that its unit ball [that of $\mathscr{R}(\hat{C}_1)$] must be linearly transformable to a polygon, which means that it should itself be a polygon. Since the unit ball of the particular example in Fig. 9 is not a polygon, we conclude that $\mathscr{R}(\hat{C}_1)$ [and consequently $\mathscr{R}(\hat{\mathcal{O}})$] is not linearly isometrically isomorphic to an $l^\infty(n)$ space for any n.

We end this section with a geometric interpretation of the approximation procedure given previously. If we apply the approximation procedure to the system in Fig. 8, the result is the system

$$\mathcal{S}_n \hat{\mathcal{O}} \mathcal{F}\left(\hat{G}_{oo}, C\right)\hat{B}_1 \mathcal{H}_n. \tag{15}$$

Looking only at the output side (the input side can be interpreted similarly using adjoints), the norm on the output side is essentially measured by sampling the elements in $\mathscr{R}(\hat{\mathcal{O}})$, that is, the norm of a function $f \in \mathscr{R}(\hat{\mathcal{O}})$ is computed by taking the $l^\infty(n)$ norm of n samples. As before, we can plot the unit ball of $\mathscr{R}(\hat{C}_1)$ in this new norm which we will call the "samples norm." (Actually, we will plot the coefficients α_1, α_2, hence the plot is two dimensional). This norm approximates the actual norm on $\mathscr{R}(\hat{C}_1)$ for large n. This approximation can be seen in Fig. 10 (for $n = 3$), where the samples norm unit ball is superimposed over the actual unit ball of $\mathscr{R}(\hat{C}_1)$. It is interesting to see that what is being done, is approximation of the unit ball of $\mathscr{R}(\hat{\mathcal{O}})$ by polygons. Thus the approximation procedure for solving the sampled-data problem can be interpreted as an approximation of norms of the input and output spaces. It is interesting to note here that the unit balls of $\mathscr{R}(\hat{\mathcal{O}})$ and $(\mathscr{R}(^*\hat{B}_1))^*$, generally represent nonlinear constraints, very much as in the continuous-time L^1 problem [6], while in discrete-time l^1 problems, the constraints are always linear. Therefore, the fact that the norms in the sampled-data problem represent nonlinear constraints (roughly speaking), seems to be a consequence of the continuous-time nature of the problem (just as in the L^1 problem). However, by essentially

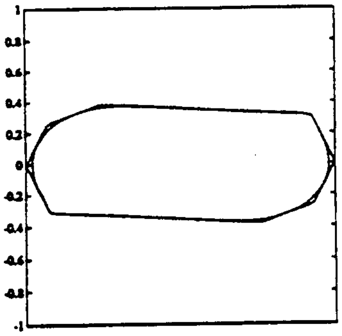


Fig. 10. The unit balls of $\mathscr{R}(\hat{C}_1)$ with the actual, and the samples norms.

approximating the nonlinear constraints by linear ones, we are able to reduce the problem to a standard discrete-time l^1 problem.

Finally, we point out that the mathematical reason behind the difference in the reductions of the $\mathscr{H}^\infty$ and l^1 sampled-data problems, is that in the former, any finite-dimensional Hilbert space is linearly isometric to $l^2(n)$, while in the latter, not every finite-dimensional Banach space is linearly isometric to $l^\infty(n)$. This reflects the fact that the isometric class of Banach spaces of dimension n is a much richer class (there is an infinite number of them, for example $l^p(n)$ for $1 \leq p \leq \infty$), than the class of Hilbert spaces of dimension n [of which there is only one, $l^2(n)$].

VII. Conclusions

This paper provides a solution for the sampled-data l^1 problem through approximation. Utilizing lifting techniques, the input/output map is decomposed in such a way that the infinite-dimensional part of the system is isolated independently of the controller. This part is then approximated in a precise way by a finite-dimensional system, whose dimension can be determined given any degree of accuracy. Computable bounds on the norm of the difference of the actual system and the approximated system are furnished, and they all depend entirely on the system's data. It is shown that the rate of convergence of this approximation is $(1/n)$.

It is interesting to note that the same approach and approximation arguments in this paper can be followed to obtain bounds like the main inequality for the L^1-induced norm sampled-data problem. A combination of this with the Riesz–Thorin convexity theorem would then show that the main inequality (with different constants) holds for general L^p-induced norm problems. In particular this holds for the L^2-induced norm case. In this case, this approximation procedure was shown to converge in [15]. The results of this paper and the above convexity argument indicate that stronger convergence at the $(1/n)$ rate actually holds. However, for the case of the L^2-induced norm sampled-data problem, an exact equivalence to a discrete-time problem can be obtained [1]. It is indicated in this paper by geometric arguments that this exact correspondence may not be possible in general for L^∞-induced norm sampled-data problems.

The approach followed in this paper is readily applicable to the structured perturbations problem for sampled-data systems [16]. The minimization problem in this set-up involves spectral radius functions, and a similar result follows from the continuity of the spectral radius function. The derivation of explicit bounds takes more work and will be reported elsewhere.

Appendix A

In the following proofs it is assumed for simplicity that the matrices D_{11} and D_{12} are zero. If D_{12} is not zero, the statement of Lemma 4 still holds. If D_{11} is not zero, the statement of Lemma 5 does not hold, however the main inequality does hold but has to be derived differently.

Proof of Lemma 4

a) If $f \in \mathscr{R}_{(^*\hat{B}_1)}$, then $f(t) = {}^*\hat{B}_1(t)x = B_1' e^{A'(\tau - t)}x$, for some $x \in \mathbb{R}^{n_x}$. We may assume without loss of generality that (A, B_1) is controllable, since if not, we can decompose the state space into the controllable and uncontrollable subspaces, and write

$$^*\hat{B}_1(t) = [B_c' \quad 0] e^{\left[\begin{smallmatrix} A_c' & 0 \\ ? & A_{nc}' \end{smallmatrix}\right](\tau - t)} T,$$

where (A_c, B_c) is controllable, T is nonsingular, and then note that $\mathscr{R}_{(^*\hat{B}_1)}$ is the same as the range of $\{B_c' e^{A_c'(\tau - t)}\}$, and thus work with (A_c, B_c) instead of (A, B_1). We also note that since the eigenvalues of A_c are a subset of the eigenvalues of A, then if τ/n is nonpathological for A, it is nonpathological for A_c.

Now, to show that $(\mathscr{H}_n \mathscr{T}_n)$ has a left inverse, we need to show that $(\mathscr{H}_n \mathscr{T}_n)$: $\mathscr{R}_{(^*\hat{B}_1)} \to L^1[0, \tau]$ is injective, but since $\mathscr{H}_n$: $l^1(n) \to L^1[0, \tau]$ is injective, it suffices to show that $\mathscr{T}_n$: $\mathscr{R}_{(^*\hat{B}_1)} \to l^1(n)$ is injective, or equivalently, that it has no null space. Given $f \in \mathscr{R}_{(^*\hat{B}_1)}$, let $\bar{f} := \mathscr{T}_n f$, since $f(t) = B_1' e^{A'(\tau - t)}x$ for some $x \in \mathbb{R}^{n_x}$, then

$$\begin{aligned} \bar{f}_i &= \frac{n}{\tau} \int_{i\tau/n}^{(i+1)\tau/n} B_1' e^{A'(\tau - t)} x \, dt \\ &= \frac{n}{\tau} B_1' \int_0^{\tau/n} e^{A'(\tau/n - \hat{t})} \, d\hat{t} \, e^{A'(n-i-1)\tau/n} x \\ &= \frac{n}{\tau} B_1' \Psi'(\tau/n) e^{A'(n-i-1)\tau/n} x, \end{aligned}$$

or in matrix notation

$$\begin{bmatrix} \bar{f}_0 \\ \vdots \\ \bar{f}_{n-1} \end{bmatrix} = \frac{n}{\tau} \begin{bmatrix} B_1' \Psi'(\tau/n) e^{A'(n-1)\tau/n} \\ \vdots \\ B_1' \Psi'(\tau/n) \end{bmatrix} x =: \frac{n}{\tau} \mathscr{B}_n' x. \quad (16)$$

Note that for $n \geq n_x$, $\mathscr{B}_n$ contains the controllability matrix of $(e^{A\tau/n}, \Psi(\tau/n)B_1)$, and since (A, B_1) is controllable and τ/n is a nonpathological sampling period, then $(e^{A\tau/n}, \Psi(\tau/n)B_1)$ is controllable, and thus the matrix $\mathscr{B}_n'$

has full rank. Therefore, if $f \in \mathscr{R}_{(^*\hat{B}_1)}$, $f \neq 0$, then $f = {}^*\hat{B}_1 x$, for some $x \in \mathbb{R}^{n_x}$, $x \neq 0$, consequently $\bar{f} \neq 0$ (since $\mathscr{B}_n'$ has full rank), implying that $\mathscr{T}_n$ has no null space and thus is injective.

To obtain the bounds we need, it is necessary to bound the norm of x that solves the equation $\bar{f} = \mathscr{B}_n' x$ by the norm of $\bar{f}$. Since $\mathscr{B}_n'$ has full rank (as a matrix), there exists a constant c_1 such that if $\bar{f} = (n/\tau)\mathscr{B}_n' x$ then $(n/\tau)\|x\|_1 \leq c_1 \|\bar{f}\|_{l^1(n)}$ (where $\|x\|_1$ is the 1-norm on $\mathbb{R}^{n_x}$). The constant c_1 can be taken as the norm of the left inverse to $\mathscr{B}_1'$. See the appendix for the proof that c_1 is independent of n.

If we define $\tilde{f} := \mathscr{H}_n \bar{f}$, then we have from the definition of $\mathscr{H}_n$: $l^1(n) \to L^1[0, \tau]$ that $\|\tilde{f}\|_{L^1[0,\tau]} = (\tau/n)\|\bar{f}\|_{l^1(n)}$. Combining this with the previous bound yields that for $\tilde{f} = (\mathscr{H}_n \mathscr{T}_n)^* \hat{B}_1 x$

$$\|x\|_1 \leq c_1 \|\tilde{f}\|_{L^1[0,\tau]}.$$

Now, to compute a bound on $\|(I - (\mathscr{H}_n\mathscr{T}_n)^{-L})|_{\mathscr{R}(\mathscr{H}_n\mathscr{T}_n {}^*\hat{B}_1)}\|$, let $\tilde{f}$ be an element in $\mathscr{R}_{(\mathscr{H}_n \mathscr{T}_n {}^*\hat{B}_1)}$, i.e., $\tilde{f} = \mathscr{H}_n \mathscr{T}_n^* \hat{B}_1 x$ for some $x \in \mathbb{R}^{n_x}$. We have already shown the existence of the left inverse $(\mathscr{H}_n\mathscr{T}_n)^{-L}$, by its definition $(\mathscr{H}_n\mathscr{T}_n)^{-L}\tilde{f} = {}^*\hat{B}_1 x$, therefore

$$\begin{aligned}
&\left\|\left(I - (\mathscr{H}_n\mathscr{T}_n)^{-L}\right)\tilde{f}\right\|_{L^1[0,\tau]} \\
&= \|\mathscr{H}_n\mathscr{T}_n^* \hat{B}_1 x - {}^*\hat{B}_1 x\|_{L^1[0,\tau]} \\
&= \int_0^\tau \left\|\left(\mathscr{H}_n\mathscr{T}_n^* \hat{B}_1 x\right)(t) - \left({}^*\hat{B}_1 x\right)(t)\right\| dt \\
&= \sum_{i=1}^{n-1} \int_{i\tau/n}^{(i+1)\tau/n} \left\|\left(\mathscr{H}_n\mathscr{T}_n^* \hat{B}_1 x\right)(t) - \left({}^*\hat{B}_1 x\right)(t)\right\| dt \\
&= \sum_{i=1}^{n-1} \int_{i\tau/n}^{(i+1)\tau/n} \left\|\frac{n}{\tau}\left(\int_{i\tau/n}^{(i+1)\tau/n} {}^*\hat{B}_1(s)\, ds\right) x \right. \\
&\qquad \left. - {}^*\hat{B}_1(t) x\right\| dt \\
&\leq \sum_{i=1}^{n-1} \int_{i\tau/n}^{(i+1)\tau/n} \left\|\frac{n}{\tau}\left(\int_{i\tau/n}^{(i+1)\tau/n} {}^*\hat{B}_1(s)\, ds\right)\right. \\
&\qquad \left. - {}^*\hat{B}_1(t)\right\| dt \|x\|_1 \\
&\leq \sum_{i=1}^{n-1} \frac{\tau^2}{n^2} \sup_{0 \leq t \leq \tau} \left\|\frac{d^*\hat{B}_1(t)}{dt}\right\| \|x\|_1 \\
&\leq \frac{\tau^2}{2n^2} \|x\|_1 \sum_{i=1}^{n-1} \sup_{0 \leq t \leq \tau} \|B_1' A' e^{A'(\tau - t)}\| \\
&\leq \frac{\tau^2}{2n^2} c_1 \|\tilde{f}\| n \|B_1'\| \, \|A'\| e^{\|A'\|\tau} \\
&\leq \frac{\tau^2}{2} c_1 \|B_1'\| \, \|A'\| e^{\|A'\|\tau} \frac{1}{n} \|\tilde{f}\|
\end{aligned} \tag{17}$$

see (18) which means that

$$\left\|\left(I - (\mathscr{H}_n\mathscr{T}_n)^{-L}\right)\right\|_{\mathscr{R}(\mathscr{H}_n\mathscr{T}_n {}^*\hat{B}_1)}\Big| \leq \frac{\tau^2}{2} c_1 \|B_1'\| \, \|A'\| e^{\|A'\|\tau} \frac{1}{n} =: \frac{K_{\hat{B}}}{n}.$$

Proof of b): By definition, $\hat{\mathscr{O}} := [\hat{C}_1 \;\; \tilde{D}_{12}]$, and

$$\begin{aligned}
\hat{\mathscr{O}}(t) &= \left[\hat{C}_1(t) \quad \tilde{D}_{12}(t)\right] = \left[C_1 e^{At} \quad C_1\left(\int_0^t e^{As}\, ds\right) B_2\right] \\
&= [0 \quad C_1] e^{\begin{bmatrix} 0 & 0 \\ I & A \end{bmatrix} t} \begin{bmatrix} 0 & B_2 \\ I & 0 \end{bmatrix},
\end{aligned}$$

where the last equality is a consequence of the formula $\int_0^t e^{As}\, ds = [0 \;\; I] e^{\begin{bmatrix} 0 & 0 \\ I & A \end{bmatrix} t} \begin{bmatrix} I \\ 0 \end{bmatrix}$. With an argument similar to that in the proof of part a), we can replace $[0 \;\; C_1]$ and $\begin{bmatrix} 0 & 0 \\ I & A \end{bmatrix}$ by C_o and A_o such that (C_o, A_o) is observable, i.e.,

$$\begin{aligned}
[0 \quad C_1] e^{\begin{bmatrix} 0 & 0 \\ I & A \end{bmatrix} t} \begin{bmatrix} 0 & B_2 \\ I & 0 \end{bmatrix} &= [C_0 \quad 0] e^{\begin{bmatrix} A_o & 0 \\ ? & A_{no} \end{bmatrix} t} T \begin{bmatrix} 0 & B_2 \\ I & 0 \end{bmatrix} \\
&= \left[C_o e^{A_o t} \quad 0\right] \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = C_0 e^{A_o t} R_1,
\end{aligned}$$

where $\begin{bmatrix} R_1 \\ R_2 \end{bmatrix} := T \begin{bmatrix} 0 & B_2 \\ I & 0 \end{bmatrix}$. Furthermore, we can replace R_1 by B_f, which is made up of the linearly independent columns of R_1, and define $\hat{\mathscr{O}}_o(t) := C_o e^{A_o t} B_f$, we then have

$$\begin{aligned}
\mathscr{R}(\hat{\mathscr{O}}) &= \mathscr{R}\left(\left\{[0 \quad C_1] e^{\begin{bmatrix} 0 & 0 \\ I & A \end{bmatrix} t} \begin{bmatrix} 0 & B_2 \\ I & 0 \end{bmatrix}\right\}\right) \\
&= \mathscr{R}\left(\left\{C_o e^{A_o t} B_f\right\}\right) = \mathscr{R}\left(\hat{\mathscr{O}}_o\right).
\end{aligned}$$

Now, to show the existence of $(\mathscr{H}_n\mathscr{S}_n)^{-L}$ on $\mathscr{R}_{(\mathscr{H}_n\mathscr{S}_n\hat{\mathscr{O}})}$, or equivalently, that $(\mathscr{H}_n\mathscr{S}_n)$ is injective, it suffices to show that $\mathscr{S}_n$ has no null space in $\mathscr{R}_{(\hat{\mathscr{O}})}$ (since $\mathscr{H}_n$: $l^\infty(n) \to L^\infty[0, \tau]$ is injective). By the representation above, if $f \in \mathscr{R}_{(\hat{\mathscr{O}})}$, $f \neq 0$, then $f(t) = C_o e^{A_o t} B_f x$ for some $x \neq 0$, $x \in \mathbb{R}^p$ (where $p \leq n_x + n_u$). Let $\bar{f} := \mathscr{S}_n f$, then $\bar{f}_i = C_o e^{A_o i\tau/n} B_f x$, or in matrix notation

$$\begin{bmatrix} \bar{f}_0 \\ \vdots \\ \bar{f}_{n-1} \end{bmatrix} = \begin{bmatrix} C_o \\ \vdots \\ C_o\left(e^{A_o \tau/n}\right)^{n-1} \end{bmatrix} B_f x =: \mathscr{C}_n B_f x.$$

Since (C_o, A_o) is observable and τ/n is not pathological, then $(C_o, e^{A_o \tau/n})$ is observable implying that the matrix $\mathscr{C}_n$ has full column rank (for $n \geq 2n_x$), and since B_f also has column rank, then $\bar{f} \neq 0$, which shows that $\mathscr{S}_n$ has no null space in $\mathscr{R}_{(\hat{\mathscr{O}})}$.

To obtain the bounds we need, it is necessary to have a bound on the norm $\|x\|_\infty$ ($\|\cdot\|_\infty$ is the maximum component norm in $\mathbb{R}^p$) of solutions of the equation $\bar{f} = \mathscr{C}_n B_f x$. Since both $\mathscr{C}_n$ and B_f have full column rank, they both have left inverses $\mathscr{C}_n^{-L}$, B_f^{-L}, and

$$\|x\|_\infty \leq \|B_f^{-L}\| \, \|\mathscr{C}_n^{-L}\| \, \|\bar{f}\|_{l^\infty(n)}.$$

Since $\mathscr{H}_n$: $l^\infty(n) \to L^\infty[0,\tau]$ preserves norms, that is, for $\bar{f} := \mathscr{H}_n \bar{f} = (\mathscr{H}_n \mathscr{S}_n)f$, we have that $\|\bar{f}\|_{L^\infty[0,\tau]} = \|f\|_{l^\infty(n)}$, the above bound becomes

$$\|x\|_\infty \le c_2 \|\bar{f}\|_{L^\infty[0,\tau]}.$$

The proof that the bound c_2 is independent of n, though long, is entirely similar to that for c_1 in part a).

Now let $f \in \mathscr{R}_{(\hat{\mathscr{O}})}$, therefore $f \in \mathscr{R}_{(\hat{\mathscr{O}}_o)}$ which means that $f = \hat{\mathscr{O}}_o x$ for some $x \in \mathbb{R}^p$. Let $\bar{f} := (\mathscr{H}_n \mathscr{S}_n)f$, by the definition of $(\mathscr{H}_n \mathscr{S}_n)^{-L}$, we have that $(\mathscr{H}_n \mathscr{S}_n)^{-L} f = f = \hat{\mathscr{O}}_o x$. We now compute,

$$\begin{aligned}
&\left\|\left(I - (\mathscr{H}_n \mathscr{S}_n)^{-L}\right)\bar{f}\right\|_{L^\infty[0,\tau]} \\
&= \sup_{0 \le t \le \tau} \left\|\left(\mathscr{H}_n \mathscr{S}_n \hat{\mathscr{O}}_o x\right)(t) - \left(\hat{\mathscr{O}}_o x\right)(t)\right\| \\
&= \sup_{0 \le i \le n-1} \sup_{0 \le \hat{t} \le \tau/n} \left\|\left(\mathscr{H}_n \mathscr{S}_n \hat{\mathscr{O}}_o x\right)(\hat{t} + i\tau/n) - \left(\hat{\mathscr{O}}_o x\right)(\hat{t} + i\tau/n)\right\| \\
&= \sup_{0 \le i \le n-1} \sup_{0 \le \hat{t} \le \tau/n} \left\|\left(\hat{\mathscr{O}}_o x\right)(i\tau/n) - \left(\hat{\mathscr{O}}_o x\right)(\hat{t} + i\tau/n)\right\| \\
&\le \sup_{0 \le i \le n-1} \sup_{0 \le \hat{t} \le \tau/n} \left\|\int_{i\tau/n}^{\hat{t}+i\tau/n} \frac{d\hat{\mathscr{O}}_o}{ds}(s)\, ds\right\| \|x\|_\infty \\
&\le \sup_{0 \le i \le n-1} \sup_{0 \le \hat{t} \le \tau/n} \int_{i\tau/n}^{\hat{t}+i\tau/n} \left\|\frac{d\hat{\mathscr{O}}_o}{ds}(s)\right\| ds \|x\|_\infty \\
&\le \sup_{0 \le i \le n-1} \int_{i\tau/n}^{(i+1)\tau/n} \left\|\frac{d\hat{\mathscr{O}}_o}{ds}(s)\right\| ds \|x\|_\infty \\
&\le \sup_{0 \le i \le n-1} \sup_{0 \le s \le \tau} \left\|\frac{d\hat{\mathscr{O}}_o}{ds}(s)\right\| \frac{\tau}{n} \|x\|_\infty \\
&\le \sup_{0 \le i \le n-1} \sup_{0 \le s \le \tau} \|C_o\| \, \|A_o\| e^{\|A_o\| s} \|B_f\| \frac{\tau}{n} \|x\|_\infty \\
&\le \|C_o\| \, \|A_o\| e^{\|A_o\|\tau} \|B_f\| \frac{\tau}{n} c_2 \|\bar{f}\|_{L^\infty[0,\tau]},
\end{aligned}$$

which results in

$$\left\|\left(I - (\mathscr{H}_n \mathscr{S}_n)^{-L}\right)\big|_{\mathscr{R}(\mathscr{H}_n \mathscr{S}_n \hat{\mathscr{O}})}\right\| \le \|C_o\| \, \|A_o\| e^{\|A_o\|\tau} \|B_f\| c_2 \tau \frac{1}{n} =: \frac{K_{\hat{\mathscr{O}}}}{n}. \quad \blacksquare$$

Proof of Lemma 5

If $\hat{D}_{11}$ comes from the lifting of a MIMO G_{11}, then $\hat{D}_{11}$ operates on vector signals, i.e., $\hat{D}_{11}$: $L_n^\infty[0,\tau] \to L_m^\infty[0,\tau]$. The induced norm of such an operator is bounded above by the maximum row sum of the matrix of the $L^\infty[0,\tau]$-induced norms of the SISO subsystems. We will prove the lemma as if $\hat{D}_{11}$ is scalar, the MIMO statement follows from the fact that if each entry in the matrix of norms tends to 0 separately, then the maximum row sum will also tend to zero.

The $L^\infty[0,\tau]$-induced norm of an operator $\mathscr{A}$ given by a kernel function $\mathscr{A}(t,s)$ is

$$\|\mathscr{A}\| = \sup_{0 \le t \le \tau} \int_0^\tau |\mathscr{A}(t,s)|\, ds.$$

The kernel function of $\hat{D}_{11}$ is given from (1) by

$$\hat{D}_{11}(t,s) = C_1 e^{A(t-s)} \mathbf{1}_{(t-s)} B_1.$$

The operator $\overline{D}_n := (\mathscr{H}_n \mathscr{S}_n)\hat{D}_{11}(\mathscr{H}_n \mathscr{T}_n)$ has a kernel function which is piecewise constant over squares of width τ/n in $[0,\tau] \times [0,\tau]$, in particular, for $t = \hat{t} + i\tau/n$ and $s = \hat{s} + j\tau/n, \hat{t}, \hat{s} \in [0, \tau/n]$

$$\overline{D}_n(t,s) = \frac{n}{\tau} C_1 e^{Ai\tau/n} \left(\int_{j\tau/n}^{(j+1)\tau/n} e^{-Ar}\, dr\right) \mathbf{1}_{(i-j-1)} B_1,$$

where $\mathbf{1}_{(\cdot)}$ is the unit step function with a discrete parameter. We now compute

$$\begin{aligned}
&\|\hat{D}_{11} - \overline{D}_n\| \\
&= \sup_{0 \le t \le \tau} \int_0^\tau \left|\hat{D}_{11}(t,s) - \overline{D}_n(t,s)\right| ds \\
&= \sup_{0 \le i \le n-1} \sup_{0 \le t \le \tau/n} \sum_{j=0}^{n-1} \int_{j\tau/n}^{(j+1)\tau/n} \cdot \left|\hat{D}_{11}(t,s) - \overline{D}_n(t,s)\right| ds \\
&= \sup_i \sup_{\hat{t}} \sum_{j=0}^{n-1} \int_{j(\tau/n)}^{(j+1)\tau/n} \cdot \left| C_1 \left(e^{A(i(\tau/n)+\hat{t}-s)} \mathbf{1}_{(t-s)} - \frac{n}{\tau} e^{Ai(\tau/n)} \cdot \int_{j(\tau/n)}^{(j+1)\tau/n} e^{-Ar}\, dr \mathbf{1}_{(i-j-1)} \right) B_1 \right| ds \\
&\le \|C_1\| \, \|B_1\| \sup_i e^{\|A\| i(\tau/n)} \sup_{\hat{t}} \sum_{j=0}^{n-1} \int_{j(\tau/n)}^{(j+1)\tau/n} \cdot \left\| e^{A(\hat{t}-s)} \mathbf{1}_{(t-s)} - \frac{n}{\tau} \cdot \int_{j(\tau/n)}^{(j+1)\tau/n} e^{-Ar}\, dr \mathbf{1}_{(i-j-1)} \right\| ds \\
&\le \|C_1\| \, \|B_1\| e^{\|A\|\tau} \sup_i \sup_{\hat{t}} \cdot \left\{ \sum_{j=0}^{i-1} \int_{j(\tau/n)}^{(j+1)\tau/n} \left\| e^{A(\hat{t}-s)} - \frac{n}{\tau} \cdot \int_{j(\tau/n)}^{(j+1)\tau/n} e^{-Ar}\, dr \right\| ds + \int_0^{\tau/n} \| e^{A(\hat{t}-s)} \|\, d\hat{s} \right\},
\end{aligned}$$

where the last term represents the case $i = j$. From (20) in Appendix B, we can bound

$$\int_{j(\tau/n)}^{(j+1)\tau/n} \| e^{A(\hat{i}-s)} - \frac{n}{\tau}\int_{j(\tau/n)}^{(j+1)\tau/n} e^{-Ar}\,dr \|\,ds$$

$$\leq \frac{\tau}{n}\| e^{A(\hat{i}-j(\tau/n))} - e^{-Aj(\tau/n)}\|$$

$$+ \frac{1}{2}\left[\sup_{j(\tau/n)\leq s \leq (j+1)\tau/n} \| Ae^{A(\hat{i}-s)}\| \right.$$

$$\left. + \sup_{j(\tau/n)\leq r \leq (j+1)\tau/n} \| Ae^{-Ar}\| \right] \frac{\tau^2}{n^2}$$

$$\leq \frac{\tau}{n} e^{\|A\| j(\tau/n)} \| e^{A\hat{i}} - I\| + \|A\| e^{\|A\|\tau} \frac{\tau^2}{n^2}$$

$$\leq e^{\|A\|\tau}\left(\frac{\tau}{n}(e^{\|A\|\hat{i}} - 1) + \|A\|\frac{\tau^2}{n^2}\right).$$

Substituting back yields

$$\|\hat{D}_{11} - \overline{D}_n\| \leq \|C_1\|\,\|B_1\| e^{2\|A\|\tau} \sup_i \sup_{\hat{i}}$$

$$\cdot \left\{ \sum_{j=0}^{i-1} \left(\frac{\tau}{n}(e^{\|A\|\hat{i}} - 1) + \|A\|\frac{\tau^2}{n^2}\right) + \frac{\tau}{n}\right\}$$

$$\leq \|C_1\|\,\|B_1\| e^{2\|A\|\tau}$$

$$\cdot \left\{ \frac{\tau}{n}(e^{\|A\|\tau} - 1) + \|A\|\frac{\tau^2}{n} + \frac{\tau}{n}\right\}$$

$$=: \frac{K_{\hat{D}}}{n}$$ ■

APPENDIX B

Integral Inequalities

Let $F(t), F_1(t), F_2(t)$ be differentiable matrix valued functions. Some useful bounds shown below can be established by using the formula

$$F(t) = F(a) + \int_a^t \frac{dF}{dt}(s)\,ds,$$

and some manipulations involving cancelling common factors and bounding the norm of an integral by the integral of the norms. Note that in the following bounds, $\|\cdot\|$ is any matrix norm provided that the same norm is used on both sides of the same inequality,

$$\int_a^b \left\| \frac{1}{b-a}\left(\int_a^b F(s)\,ds\right) - F(t)\right\| dt \leq \frac{(b-a)^2}{2}\left(\sup_{a\leq t\leq b} \left\|\frac{dF}{dt}(t)\right\|\right) \tag{18}$$

$$\left\| \int_0^T F(t)F'(t)\,dt - \frac{1}{T}\left(\int_0^T F(s)\,ds\right)\left(\int_0^T F'(r)\,dr\right)\right\| \leq 2T^3\left(\sup_{0\leq t\leq T}\left\|\frac{dF}{dt}\right\|\right)^2 \tag{19}$$

$$\int_a^b \left\| F_1(t) - \frac{1}{b-a}\left(\int_a^b F_2(r)\,dr\right)\right\| dt \leq |b-a|\,|F_1(a) - F_2(a)| + \frac{1}{2}\left(\sup_{a\leq t\leq b}\left\|\frac{dF_1}{dt}(t)\right\| + \sup_{a\leq t\leq b}\left\|\frac{dF_2}{dt}(t)\right\|\right) \cdot |b-a|^2. \tag{20}$$

Completion of Proof of Lemma 4-a)

Claim: c_1 is independent of n.

Proof: We will construct c_1 as an upper bound on the norm of the left inverse to $\mathscr{B}'_n$. This is done by taking the pseudo-inverse as a left inverse to $\mathscr{B}'_n$, and finding a bound on its norm that is independent of n. The pseudo-inverse to $\mathscr{B}'_n$ is $(\mathscr{B}_n\mathscr{B}'_n)^{-1}\mathscr{B}_n$, and note that the inverse exists since $\mathscr{B}'_n$ has full column rank. We first bound $\|(\mathscr{B}_n\mathscr{B}'_n)^{-1}\|$. From the definition of $\mathscr{B}'_n$, we have

$$\mathscr{B}_n\mathscr{B}'_n = \sum_{i=0}^{n-1} e^{Ai\tau/n}\Psi(\tau/n)B_1B_1'\Psi'(\tau/n)e^{A'i\tau/n}.$$

Denote the controllability Grammian over the finite time τ, by

$$W_\tau := \int_0^\tau e^{At}B_1B_1'e^{A't}\,dt.$$

We will first show that $(n/\tau)(\mathscr{B}_n\mathscr{B}_n')\overset{n\to\infty}{\to} W_\tau$

$$\begin{aligned}
&\|W_\tau - \frac{n}{\tau}(\mathscr{B}_n\mathscr{B}_n')\|\\
&= \|\int_0^\tau e^{At}B_1B_1'e^{A't}\,dt - \frac{n}{\tau}\sum_{i=0}^{n-1} e^{Ai\tau/n}\Psi(\tau/n)B_1B_1'\Psi'(\tau/n)e^{A'i\tau/n}\|\\
&= \|\sum_{i=0}^{n-1}\int_{i\tau/n}^{(i+1)\tau/n} e^{At}B_1B_1'e^{A't}\,dt - \frac{n}{\tau}\sum_{i=0}^{n-1} e^{Ai\tau/n}\Psi(\tau/n)B_1B_1'\Psi'(\tau/n)e^{A'i\tau/n}\|\\
&\le \sum_{i=0}^{n-1}\|e^{Ai\tau/n}\left(\int_0^{\tau/n} e^{A\hat{t}}B_1B_1'e^{A'\hat{t}}\,d\hat{t}\right)e^{A'i\tau/n} - \frac{n}{\tau}e^{Ai\tau/n}\Psi(\tau/n)B_1B_1'\Psi'(\tau/n)e^{A'i\tau/n}\|\\
&\le \sum_{i=0}^{n-1} e^{2\|A\|i\tau/n}\|\int_0^{\tau/n} e^{A\hat{t}}B_1B_1'e^{A'\hat{t}}\,d\hat{t} - \frac{n}{\tau}\left(\int_0^{\tau/n}e^{As}\,ds\right)B_1B_1'\left(\int_0^{\tau/n}e^{A'r}\,dr\right)\|\\
&\le \sum_{i=0}^{n-1} e^{2\|A\|i\tau/n}2\frac{\tau^3}{n^3}\cdot\left(\sup_{0\le\hat{t}\le\tau/n}\|B_1\|\,\|A\|e^{\|A\|\hat{t}}\right)^2 \qquad (21)
\end{aligned}$$

where the last step is a consequence of formula (19). After bounding $e^{2\|A\|i\tau/n}\le e^{2\|A\|\tau}$ and summing to yield a factor of n, (21) becomes

$$\|W_\tau - \frac{n}{\tau}(\mathscr{B}_n\mathscr{B}_n')\| \le 2\,e^{4\|A\|\tau}\|B_1\|^2\|A\|^2\frac{\tau^3}{n^2} =: M_1\frac{1}{n^2},$$

where M_1 is a constant. Now, since $(n/\tau)(\mathscr{B}_n\mathscr{B}_n')\overset{n\to\infty}{\to} W_\tau$, it follows that $((n/\tau)\mathscr{B}_n\mathscr{B}_n')^{-1}\overset{n\to\infty}{\to} W_\tau^{-1}$ [18, theorem 10.12]. An explicit bound (for large n) on the norm of $((n/\tau)\mathscr{B}_n\mathscr{B}_n')^{-1}$ in terms of the norm of W_τ^{-1} can be constructed in several ways, one way is by [18, theorem 10.11]

$$\begin{aligned}
\|\left(\frac{n}{\tau}\mathscr{B}_n\mathscr{B}_n'\right)^{-1}\| &\le \|W_\tau^{-1}\| + \|W_\tau^{-1}\|^2\|W_\tau - \left(\frac{n}{\tau}\mathscr{B}_n\mathscr{B}_n'\right)\|\\
&\quad + 2\|W_\tau^{-1}\|^3\|W_\tau - \left(\frac{n}{\tau}\mathscr{B}_n\mathscr{B}_n'\right)\|^2\\
&\le \|W_\tau^{-1}\| + \|W_\tau^{-1}\|^2 M_1 + 2\|W_\tau^{-1}\|^3 M_1^2\\
&=: M_2,
\end{aligned}$$

for $n^2 > 2M_1\|W_\tau^{-1}\|$. To take care of the case of n such that $n^2\le 2M_1\|W_\tau^{-1}\|$, note that is only a finite number of such n's, and let M_3 be the maximum of $\|((n/\tau)\mathscr{B}_n\mathscr{B}_n')^{-1}\|$ over this finite set of n's (note also that $\|(\mathscr{B}_n\mathscr{B}_n')^{-1}\|$ exists if $n\ge n_x$ and τ/n is not a pathological sampling period). Letting $M_4 := \max\{M_2, M_3\}$, we obtain

$$\|\left(\frac{n}{\tau}\mathscr{B}_n\mathscr{B}_n'\right)^{-1}\| \le M_4 \Rightarrow \|(\mathscr{B}_n\mathscr{B}_n')^{-1}\| \le \frac{n}{\tau}M_4$$

$\forall n\ge n_x$ such that τ/n is not pathological. Finally, to find $\|(\mathscr{B}_n\mathscr{B}_n')^{-1}\mathscr{B}_n\|$, note that this is the induced norm from $l^1(n)$ to $\mathbb{R}^{n_x}$ with the $\|\cdot\|_1$ norm, i.e., it is the maximum column sum norm on the matrix, therefore

$$\begin{aligned}
\|(\mathscr{B}_n\mathscr{B}_n')^{-1}\mathscr{B}_n\| &\le \|(\mathscr{B}_n\mathscr{B}_n')^{-1}\|\cdot\|\Psi(\tau/n)\left[(e^{A\tau/n})^{n-1}B_1\ \cdots\ B_1\right]\|\\
&\le \|(\mathscr{B}_n\mathscr{B}_n')^{-1}\|\|\Psi(\tau/n)\|\max\cdot\{\|(e^{A\tau/n})^{n-1}\|,\cdots,\|e^{A\tau/n}\|\}\|B_1\|\\
&\le \frac{n}{\tau}M_4\frac{\tau}{n}e^{\|A\|\tau/n}e^{\|A\|\tau}\|B_1\|\\
&\le M_4\,e^{2\|A\|\tau}\|B_1\| =: c_1,
\end{aligned}$$

since $\|\Psi(\tau/n)\| = \|\int_0^{\tau/n}e^{As}\,ds\| \le \int_0^{\tau/n}e^{\|A\|s}\,ds \le \int_0^{\tau/n}e^{\|A\|\tau/n}\,ds \le (\tau/n)e^{\|A\|\tau/n}$. This yields the desired bound c_1 which is independent of n. □

Existence of Preadjoints

Given an operator $H: X^*\to X^*$, where X^* is the dual of some Banach space X, its preadjoint ${}^*H$ is such that ${}^*H: X\to X$ and $({}^*H)^* = H$. Not every operator has a preadjoint, but the operators that we are dealing with do. For example, $\hat{B}_1: L^\infty[0,\tau]\to\mathbb{R}^{n_x}$ has a preadjoint ${}^*\hat{B}_1: \mathbb{R}^{n_x}\to L^1[0,\tau]$. Let $\hat{B}_1(s)$ denote the matrix valued kernel function representing the operator $\hat{B}_1$, it is very easy to check that the operator from $\mathbb{R}^{n_x}$ to $L^1[0,\tau]$ given by the matrix valued kernel function $\hat{B}_1'(t)$ (here $'$ denotes matrix transpose) is a preadjoint to the operator $\hat{B}_1$.

REFERENCES

[1] B. Bamieh and J. B. Pearson, "A general framework for linear periodic systems with application to $\mathscr{H}^\infty$ sampled-data control," *IEEE Trans. Automat. Contr.*, vol. 37, no. 4, pp. 418–435, Apr. 1992.

[2] B. Bamieh, J. B. Pearson, B. A. Francis, A. Tannenbaum, "A lifting technique for linear periodic systems with applications to sampled-data control," *Syst. Contr. Lett.*, vol. 17, pp. 79–88, 1991.

[3] B. Bamieh, "Analysis and robust control of Hybrid continuous/discrete-time systems," Ph.D. dissertation, Rice University, July 1991.

[4] T. Chen and B. Francis, "On the L^2 induced norm of a sampled-data system," *Syst. Control Lett.*, vol. 15, pp. 211–219, 1990.

[5] M. A. Dahleh and J. B. Pearson, "l^1-optimal feedback control for MIMO discrete-time systems," *IEEE Trans. Automat. Contr.*, vol. AC-32, 1987.

[6] ——, "L^{-1}-optimal compensators for continuous-time systems," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 889–895, 1987.

[7] M. A. Dahleh, P. Voulgaris and L. Valavani, "Optimal and robust controllers for periodic and multirate systems," to appear in *IEEE Trans. Automat. Contr.*

[8] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. New York: Academic, 1975.

[9] I. J. Diaz-Bobillo and M. A. Dahleh, "Minimization of the maximum peak-to-peak gain: the multiblock problem," submitted to the *IEEE Trans. Automat. Contr.*, 1992.

[10] G. Dullerud and B. Francis, "$\mathscr{L}^1$ analysis and design of sampled-data systems," *IEEE Trans. Automat. Contr.*, vol. 37, no. 4, pp. 436–446, Apr. 1992.

[11] N. Dunford and J. T. Schwartz, *Linear Operators*. New York: Interscience, 1958.

[12] B. A. Francis and T. T. Georgiou, "Stability theory for linear time-invariant plants with periodic digital controllers," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 820–832, 1988.

[13] P. T. Kabamba and S. Hara, "Worst case analysis and design of sampled data control systems," preprint.

[14] J. P. Keller and B. D. O. Anderson, "A new approach to discretization of continuous-time controllers," to appear in *Automatica*.

[15] ——, "A new approach to the discretization of continuous-time controllers," *IEEE Trans. Automat. Contr.*, vol. 37, no. 2, pp. 214–223, Feb. 1992.

[16] M. H. Khammash, "Necessary and sufficient conditions for the robustness of time-varying systems with applications to sampled-data systems," preprint, submitted to the *IEEE Trans. Automat. Contr.*.

[17] J. S. McDonald and J. B. Pearson, "l^1-Optimal control of multivariable systems with output norm constraints," *Automatica*, vol. 27, no. 3, 1991.

[18] W. Rudin, *Functional Analysis*. New York: McGraw-Hill, 1973.

[19] N. Sivashankar and Pramod P. Khargonekar, "$\mathscr{L}_\infty$-induced norm of sampled-data systems," in *Proc. CDC*, 1991.

[20] A. E. Taylor and D. C. Lay, *Introduction to Functional Analysis*, 2nd ed. New York: Wiley, 1980.

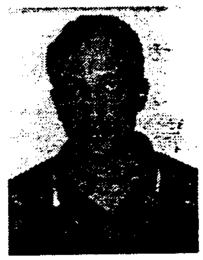
[21] P. M. Thompson, G. Stein, and M. Athans, "Conic sectors for sampled-data feedback systems," *Syst. Contr. Lett.*, vol. 3, pp. 77–82, 1983.

[22] H. T. Toivonen, "Sampled-data control of continuous-time systems with an $\mathscr{H}_\infty$ optimality criterion," *Automatica*, vol. 28, no. 1, pp. 45–54, 1992.

[23] Y. Yamamoto, "A new approach to sampled-data control systems-a function space viewpoint with applications to tracking problems," in *Proc. 29th CDC*, 1990, pp. 1882–1887.

Bassam A. Bamieh (S'87–M'91) was born in Jordan in 1963. He received the B.S. degree in electrical engineering from Valparaiso University, Valparaiso, IN, in 1983, and the M.Sc. and Ph.D. degrees from Rice University, Houston, TX, in 1986 and 1992, respectively.

He is currently an Assistant Professor in the Department of Electrical and Computer Engineering and in the Coordinated Science Laboratory at the University of Illinois, Urbana-Champaign. His general area of interest is control and system theory; his current areas of interest include robust control, infinite-dimensional systems, and digital control and sampled-data systems.



Munther A. Dahleh (S'84–M'87) was born in August 1962. He received the B.S.E.E. degree from Texas A & M University, College Station, TX in 1983, and the Ph.D. degree from Rice University, Houston, TX 1987.

Since 1987, he has been with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, where he is currently an Associate Professor. In July 1992 he was awarded the Finmeccanica career development chair. He has held consulting positions with NASA and the C.S. Draper Laboratory since 1988, and with FIAT Research Center since 1991. His current interests include robust control, identification of uncertain systems and adaptation, and control of time-varying and infinite-dimensional systems.

Dr. Dahleh is currently serving as an Associate Editor for the *Systems and Control Letters*. He is a receipient of the 1993 Eckman award.

J. Boyd Pearson, Jr., (S'57–M'58–M'65–SM'77–F'79) received the Ph.D. degree in 1962 from Purdue University, Lafayette, IN.

He served as Assistant Professor of Electrical Engineering at Purdue from 1962 to 1965, when he joined Rice University, Houston, TX, where he is now a J.S. Abercrombie Professor.

Dr. Pearson served as President of the IEEE Control Systems Society in 1984 and received outstanding paper awards in 1985 with Bor-Chin Chang and in 1989 with Munther Dahleh.

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE

3. REPORT TYPE AND DATES COVERED
FINAL/15 SEP 91 TO 14 OCT 94

4. TITLE AND SUBTITLE
FUNDAMENTAL LIMITATIONS AND CAPABILITIES OF ROBUST IDENTIFICATION AND ADAPTIVE CONTROL

5. FUNDING NUMBERS
2304/AS
AFOSR-91-0368

6. AUTHOR(S)
M. A. DAHLEH

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
LABORATORY FOR INFORMATION AND DECISION SYSTEMS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MA 02139-4307

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR- 95 0101

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
110 DUNCAN AVE, SUTE B115
BOLLING AFB DC 20332-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFOSR-91-0368

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION IS UNLIMITED

13. ABSTRACT (Maximum 200 words)

In the past decade, a powerful theory for designing robust control systems has emerged. Starting with a model, and a description of the uncertainty (structured, parametric etc.), a controller can be designed to meet a variety of performance specifications. This development, however, has not been accompanied by parallel development in system identification methods by which a plant model and a description of uncertainty is provided. In an attempt to bridge this gap, a new area of research in robust identification has emerged in the last few years.

14. SUBJECT TERMS

UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR(SAME AS REPORT)